(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 May 2003 (08.05.2003)

PCT

(10) International Publication Number
WO 03/038113 A2

(51) International Patent Classification[7]: C12Q

(21) International Application Number: PCT/US02/34730

(22) International Filing Date: 29 October 2002 (29.10.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/335,004 31 October 2001 (31.10.2001) US

(71) Applicant *(for all designated States except US)*: MILLENNIUM PHARMACEUTICALS, INC. [US/US]; 75 Sidney Street, Cambridge, MA 02139 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: WILLIAMSON, Mark [US/US]; 15 Stonecrest Drive, Saugus, MA 01906 (US). RUDOLPH-OWEN, Laura, A. [US/US]; 186 Arborway, Apt. 1, Jamaica Plain, MA 02130 (US).

(74) Agent: BOSSONE, Steven, A.; Millinennium Pharmaceuticals, Inc., 75 Sidney Street, Cambridge, MA 02139 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/038113 A2

(54) Title: METHODS AND COMPOSITIONS FOR THE TREATMENT AND DIAGNOSIS OF CELLULAR PROLIFERATION DISORDERS USING 25943

(57) Abstract: The present invention relates to methods and compositions for the treatment and diagnosis of cellular proliferation disorders, including, but not limited to, breast cancer, ovarian cancer, lung cancer, and colon cancer. The invention further provides methods for identifying a compound capable of treating a cellular proliferation disorders disorder or modulating cellular proliferation. The invention also provides a method for modulating cellular proliferation e.g., modulating cellular proliferation in a subject. In addition, the invention provides a method for treating a subject having a cellular proliferation disorder characterized by aberrant 25943 polypeptide activity or aberrant 25943 nucleic acid expression.

METHODS AND COMPOSITIONS FOR THE TREATMENT AND DIAGNOSIS OF CELLULAR PROLIFERATION DISORDERS USING 25943

This application claims priority to U.S. provisional application number 60/335,004, filed October 31, 2001, the entire contents of which are herein incorporated by reference.

Colorectal cancer is the fourth most common cancer worldwide and the second most common cause of cancer deaths. Within the United States alone, there will be over 150,000 new cases and 55,000 deaths this year. In fact, it is postulated that 50% of the Western population will develop a colorectal tumor by the age of 70, with 10% of these tumors progressing to malignancy. Despite advances in therapeutic treatment, the prognosis remains poor, with only a five-year survival rate around 45%. Although the progression of the disease has been well characterized (areas of dysplasia within the colon develop into polyps, which eventually have the potential to become adenocarcinomas; adenocarcinomas become invasive and metastasize to various regions of the body, predominately the liver), diagnosis is primarily made during later stages of the disease.

Ovarian cancer, the deadliest of the gynecologic cancers, is the fifth leading cause of cancer death among U.S. women; an estimated 13,900 American women will die from ovarian cancer in 2001. Ovarian cancer occurs in one out of 55 women, and the number of women diagnosed with the disease is projected to increase slightly, from 23,100 new cases in 2000 to 23,400 expected cases in 2001. Currently, 50 percent of the women diagnosed with ovarian cancer die from it within five years. When detected early, ovarian cancer is very treatable, but the vast majority of cases are not diagnosed until the cancer has spread beyond the ovaries. For example, in cases where ovarian cancer is detected before it has spread beyond the ovaries, more than 90 percent of women will survive longer than five years. However, only 25 percent of ovarian cancer cases in the U.S. are diagnosed in the beginning stages; when diagnosed in advanced stages, the chance of five-year survival is only about 25 percent. Ovarian cancer may be difficult to diagnose because symptoms are easily confused with other diseases, and because there is no reliable, easy-to-administer screening tool.

Lung cancer is among the most common cancers in the Western world. In the United States, there were approximately 170,000 new cases of lung cancer in 1999. Since the mid-1990s, about 150,000 Americans have died each year from this disease. Lung cancer is the leading category of cancer death in men, and - since the late 1980s - it has surpassed breast cancer as the leading category of cancer death in women. Findings from

the U.S. National Cancer Institute (NCI) indicate that the upward trend in cancer-related death is due to the rapidly increasing rate of lung cancer mortality. Statistical projections suggest that lung cancer mortality in this decade will continue to rise to a rate of over 50 deaths per year per 100,000 population in America. Current lung cancer prevention programs are not expected to influence lung cancer death rates until after the year 2000. There is a close relationship between the number of lung cancer cases and lung cancer deaths in America. This is because of the low 5-year survival rate for this disease. Although lung cancer survival rates have improved over the last 40 years, the percentage (approximately 13%) continues to be low in comparison to other cancers.

Given the prevalence of these disorders, and the lack of effective cures and early diagnostics, there currently exists a great need for methods and compositions which can serve as markers before the onset of symptoms and which can serve as a means for identifying therapeutics to treat and or cure these disorders.

The present invention provides methods and compositions for the diagnosis and treatment of cellular proliferation disorders. The present invention is based, at least in part, on the discovery that expression of the 25943 gene (a glycosylasparaginase) is upregulated in tumors (*e.g.*, ovarian, lung, breast, and colon tumors). The present invention is further based, at least in part, on the discovery that 25943 expression is regulated during the cell cycle. The invention is still further based, at least in part, on the discovery that 25943 may be involved in the post-translational modification and processing of proteins. Without intending to be limited by mechanism, it is believed that modulation, *e.g.*, inhibition, of 25943 activity may modulate, *e.g.*, inhibit, protein modulation (*e.g.*, post-translational modification and process) in tumor cells that is relevant to their tumorigenic potential, therefore, modulating, *e.g.*, inhibiting cellular proliferation and tumorigenesis.

Accordingly, the present invention provides methods for the diagnosis and treatment of cellular proliferation disorders including but not limited to cancer, *e.g.*, breast cancer, ovarian cancer, lung cancer, and colon cancer.

In one aspect, the invention provides methods for identifying a compound capable of treating a cellular proliferation disorder, *e.g.*, breast cancer, ovarian cancer, lung cancer, and colon cancer. The method includes assaying the ability of the compound to modulate 25943 nucleic acid expression or 25943 polypeptide activity. In one embodiment, the ability of the compound to modulate nucleic acid expression or 25943 polypeptide activity is determined by detecting the glycosylasparaginase activity of a cell. In another

embodiment, the ability of the compound to modulate nucleic acid expression or 25943 polypeptide activity is determined by detecting modulation of cellular proliferation in a cell.

In another aspect, the invention provides methods for identifying a compound capable of modulating cellular proliferation. The method includes contacting a cell expressing a 25943 nucleic acid or polypeptide (*e.g.*, a breast cell, a breast tumor cell, an ovary cell, an ovarian tumor cell, a lung cell, a lung tumor cell, a colon cell, and/or a colon tumor cell) with a test compound and assaying the ability of the test compound to modulate the expression of a 25943 nucleic acid or the activity of a 25943 polypeptide.

In a further aspect, the invention features a method for modulating cellular proliferation. The method includes contacting a cell (*e.g.*, a breast cell, a breast tumor cell, an ovary cell, an ovarian tumor cell, a lung cell, a lung tumor cell, a colon cell, and/or a colon tumor cell) with a 25943 modulator, for example, an anti-25943 antibody, a 25943 polypeptide comprising the amino acid sequence of SEQ ID NO:2, or a fragment thereof, a 25943 polypeptide comprising an amino acid sequence which is at least 90 percent identical to the amino acid sequence of SEQ ID NO:2, an isolated naturally occurring allelic variant of a polypeptide consisting of the amino acid sequence of SEQ ID NO:2, a small molecule, an antisense 25943 nucleic acid molecule, a nucleic acid molecule of SEQ ID NO:1, or a fragment thereof, or a ribozyme.

In yet another aspect, the invention features a method for treating a subject having a cellular proliferation disorder, *e.g.*, a cellular proliferation disorder characterized by aberrant 25943 polypeptide activity or aberrant 25943 nucleic acid expression, such as breast cancer, ovarian cancer, lung cancer, and colon cancer. The method includes administering to the subject a therapeutically effective amount of a 25943 modulator, *e.g.*, in a pharmaceutically acceptable formulation or by using a gene therapy vector. In one embodiment, the 25943 modulator may be a small molecule, an anti-25943 antibody, a 25943 polypeptide comprising the amino acid sequence of SEQ ID NO:2, or a fragment thereof, a 25943 polypeptide comprising an amino acid sequence which is at least 90 percent identical to the amino acid sequence of SEQ ID NO:2, an isolated naturally occurring allelic variant of a polypeptide consisting of the amino acid sequence of SEQ ID NO:2, an antisense 25943 nucleic acid molecule, a nucleic acid molecule of SEQ ID NO:1, or a fragment thereof, or a ribozyme.

In another aspect, the invention provides a method for modulating, *e.g.*, increasing or decreasing, cellular proliferation in a subject by administering to the subject a 25943 modulator.

In another embodiment, the invention provides a 25943 nucleic acid molecule which is at least 81%, 82%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5% or more identical to the nucleotide sequence (*e.g.*, to the entire length of the nucleotide sequence) shown in SEQ ID NO:1 or 3, or a complement thereof.

In a preferred embodiment, the isolated nucleic acid molecule includes the nucleotide sequence shown SEQ ID NO:1 or 3, or a complement thereof. In another preferred embodiment, the nucleic acid molecule consists of the nucleotide sequence shown in SEQ ID NO:1 or 3. In another preferred embodiment, the nucleic acid molecule includes a fragment of at least 1045, 1046, 1047, 1050, 1075, 1100, 1200, 1300, 1350 or more nucleotides (*e.g.*, contiguous nucleotides) of the nucleotide sequence of SEQ ID NO:1 or 3, or a complement thereof.

In another embodiment, an 25943 nucleic acid molecule includes a nucleotide sequence encoding a protein having an amino acid sequence sufficiently identical to the amino acid sequence of SEQ ID NO:2. In a preferred embodiment, a 25943 nucleic acid molecule includes a nucleotide sequence encoding a protein having an amino acid sequence at least 76%, 77%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5% or more identical to the entire length of the amino acid sequence of SEQ ID NO:2.

In another preferred embodiment, an isolated nucleic acid molecule encodes the amino acid sequence of human 25943. In yet another preferred embodiment, the nucleic acid molecule includes a nucleotide sequence encoding a protein having the amino acid sequence of SEQ ID NO:2

In other preferred embodiments, the nucleic acid molecule encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the nucleic acid molecule hybridizes to a complement of a nucleic acid molecule comprising SEQ ID NO:1 or 3 under stringent conditions.

Another aspect of this invention features isolated or recombinant 25943 proteins and polypeptides. In a preferred embodiment, the 25943 protein family member has an amino acid sequence at least about 76%, 77%, 80%, 85%, 86%, 87%, 88%, 89%, 90%,

91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5% or more identical to the amino acid sequence of SEQ ID NO:2.

In yet another preferred embodiment, the 25943 protein family member is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a complement of a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3.

In another embodiment, the invention features fragments of the protein having the amino acid sequence of SEQ ID NO:2, wherein the fragment comprises at least 232, 233, 234, 250, 275, or 300 amino acids (*e.g.*, contiguous amino acids) of the amino acid sequence of SEQ ID NO:2. In another embodiment, the protein, preferably a 25943 protein, has the amino acid sequence of SEQ ID NO:2.

In another embodiment, the invention features an isolated 25943 protein family member which is encoded by a nucleic acid molecule consisting of a nucleotide sequence at least about 81%, 82%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5% or more identical to a nucleotide sequence of SEQ ID NO:1 or 3, or a complement thereof. This invention further features an isolated protein, preferably a 25943 protein, which is encoded by a nucleic acid molecule consisting of a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising a complement of a the nucleotide sequence of SEQ ID NO:1 or 3.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

Table 1 depicts the expression levels of human 25943 mRNA in various human cell types and tissues, as determined by Taqman analysis. Sample No.: (1) normal artery; (2) diseased aorta; (3) normal vein; (4) coronary smooth muscle cells; (5) human umbilical vein endothelial cells (HUVECs); (6) normal heart; (7) heart (congestive heart failure); (8) kidney; (9) skeletal muscle; (10) normal adipose tissue; (11) pancreas; (12) primary osteoblasts; (13) differentiated osteoclasts; (14) normal skin; (15) normal spinal cord; (16) normal brain cortex; (17) normal brain hypothalamus; (18) nerve; (19) dorsal root ganglion; (20) normal breast; (21) breast tumor; (22) normal ovary; (23) ovarian tumor; (24) normal prostate; (25) prostate tumor; (26) salivary gland; (27) normal colon; (28) colon tumor; (29) lung tumor; (30) lung (chronic obstructive pulmonary disease); (31) colon (inflammatory bowel disease); (32) normal liver; (33) liver fibrosis; (34) normal spleen; (35) normal tonsil; (36) normal lymph node; (37) normal small intestine; (38)

macrophages; (39) synovium; (40) activated peripheral blood mononuclear cells; (41) neutrophils; (42) megakaryocytes; (43) erythroid cells; (44) positive control.

Table 2 depicts the expression levels of human 25943 mRNA in various human tumors, as determined by Taqman analysis. Sample No.: (1-3) normal breast; (4) breast tumor (MD-IDC); (5) breast tumor; (6) breast tumor (PD-); (7) breast tumor (IDC); (8) breast tumor (ILC (LG)); (9) lymph; (10) lung (breast metastasis); (11-12) normal ovary; (13-17) ovary tumor; (18-20) normal lung; (21) lung tumor (SmC); (22-23) lung tumor (PDNSCC); (24) lung tumor (SCC); (25-26) lung tumor (ACA); (27-29) normal colon; (30-31) colon tumor (MD); (32) colon tumor; (33) colon tumor (MD-PD); (34-35) colon tumor - liver metastasis; (36) normal liver (female); (37-38) cervix - squamous cell carcinoma; (39) human microvascular endothelial cells (HMVECs) - arrested; (40) human microvascular endothelial cells (HMVECs) - proliferating; (41) hemangioma; (42) HCT116 cells - normoxic; (43) HCT116 cells - hypoxic.

Table 3 depicts the expression levels of human 25943 mRNA in various xenograft (tumorigenic) cell lines, as determined by Taqman analysis. Sample No.: (1) MCF-7 breast tumor; (2) ZR75 breast tumor; (3) T47D breast tumor; (4) MDA 231 breast tumor; (5) MDA 435 breast tumor; (6) SKBr3 breast tumor; (7) DLD 1 colon tumor (stage C); (8) SW480 colon tumor (stage B); (9) SW620 colon tumor (stage C); (10) HCT 116 colon tumor; (11) HT29 colon tumor; (12) colo 205 colon tumor; (13) NCIH125 lung tumor; (14) NCIH67 lung tumor; (15) NCIH322 lung tumor; (16) NCIH460 lung tumor; (17) A549 lung tumor; (18) normal human bronchial epithelium (NHBE); (19) SKOV-3 ovary tumor; (20) OVCAR-3 ovary tumor; (21) 293 baby kidney cells; (22) 293T baby kidney cells.

Table 4 depicts the expression levels of human 25943 mRNA in various xenograft (tumorigenic) cell lines, as determined by Taqman analysis. Sample No.: (1) MCF-7 breast tumor; (2) ZR75 breast tumor; (3) T47D breast tumor; (4) MDA 231 breast tumor; (5) MDA 435 breast tumor; (6) SKBr3 breast tumor; (7) DLD 1 colon tumor (stage C); (8) SW480 colon tumor (stage B); (9) SW620 colon tumor (stage C); (10) HCT 116 colon tumor; (11) HT29 colon tumor; (12) colo 205 colon tumor; (13) NCIH125 colon tumor; (14) NCIH67 colon tumor; (15) NCIH322 colon tumor; (16) NCIH460 colon tumor; (17) A549 colon tumor; (18) normal human bronchial epithelium (NHBE); (19) SKOV-3 ovary tumor; (20) OVCAR-3 ovary tumor; (21) 293 baby kidney cells; (22) 293T baby kidney cells.

Table 5 depicts the expression levels of human 25943 mRNA in various staged colon tumors, as determined by Taqman analysis. Sample No.: (1-5) normal colon; (6-7) adenomas; (8-12) colonic ACA-B; (13-18) colonic ACA-C; (19-24) normal liver; (25-30) liver metastasis.

Table 6 depicts the expression levels of human 25943 mRNA in various staged colon tumors, as determined by Taqman analysis. Sample No.: (1-5) normal colon; (6-7) adenomas; (8-12) colonic ACA-B; (13-18) colonic ACA-C; (19-24) normal liver; (25-30) liver metastasis.

Table 7 depicts the expression levels of human 25943 mRNA in various colon metastases, as determined by Taqman analysis. Sample No.: (1-3) normal colon; (4-5) colonic ACA-C; (6) colonic ACA-B; (7) adenocarcinoma; (8-24) colon metastasis to the liver; (25-27) normal liver.

Table 8 depicts the expression levels of human 25943 mRNA in a k-ras disrupted DLD-1 and HCT116 colon tumor cells, as determined by Taqman analysis. Sample No.: (1-4) DLD-1; (5-9) HCT-116 cells.

Table 9 depicts the expression levels of human 25943 mRNA in synchronized tumor cells induced to progress through the cell cycle, as determined by Taqman analysis. Sample No.: (1) HCT116, aphidicolin, t = 0; (2) HCT116, aphidicolin, t = 3; (3) HCT116, aphidicolin, t = 6; (4) HCT116, aphidicolin, t = 9; (5) HCT116, aphidicolin, t = 12; (6) HCT116, aphidicolin, t = 15; (7) HCT116, aphidicolin, t = 18; (8) HCT116, aphidicolin, t = 21; (9) HCT116, aphidicolin, t = 24; (10) HCT116, nocodazole, t = 0; (11) HCT116, nocodazole, t = 3; (12) HCT116, nocodazole, t = 6; (13) HCT116, nocodazole, t = 9; (14) HCT116, nocodazole, t = 15; (15) HCT116, nocodazole, t = 18; (16) HCT116, nocodazole, t = 21; (17) HCT116, nocodazole, t = 24; (18) DLD, nocodazole, t = 3; (19) DLD, nocodazole, t = 6; (20) DLD, nocodazole, t = 9; (21) DLD, nocodazole, t = 12; (22) DLD, nocodazole, t = 15; (23) DLD, nocodazole, t = 18; (24) DLD, nocodazole, t = 21; (25) A549, mimo, t = 0; (26) A549, mimo, t = 3; (27) A549, mimo, t = 6; (28) A549, mimo, t = 9; (29) A549, mimo, t = 15; (30) A549, mimo, t = 18; (31) A549, mimo, t = 21; (32) A549, mimo, t = 24; (33) MCF10A, mimo, t = 0; (34) MCF10A, mimo, t = 3; (35) MCF10A, mimo, t = 6; (36) MCF10A, mimo, t = 9; (37) MCF10A, mimo, t = 12; (38) MCF10A, mimo, t = 18; (39) MCF10A, mimo, t = 21; (40) MCF10A, mimo, t = 24.

Table 10 depicts the expression levels of human 25943 mRNA in various *in vitro* oncogene cell models, as determined by Taqman analysis. Sample No.: (1) SMAD4-SW480 control; (2) SMAD4-SW480 24 hours; (3) SMAD4-SW480 48 hours; (4) SMAD4-

SW480 72 hours; (5) L51747 mucinous; (6) HT29 non-mucinous; (7) SW620 non-mucinous; (8) CSC-1 normal; (9) NCM-260 normal; (10) HCT116 RER+; (11) SW48 RER+; (12) SW480 RER -/-; (13) CACO RER -/-; (14) JDLD-1; (15) JHCT116; (16) DKO1; (17) DKO4; (18) DKS-8; (19) Hke3; (20) HKh2; (21) HK2-6; (22) e3Ham#9; (23) APC5 -/-; (24) APC6 -/-; (25) APC1 +/+; (26) APC13 +/+

Table 11 depicts the expression levels of human 25943 mRNA in various ovarian cell tumor models, as determined by Taqman analysis. Sample No.: (1) SKOV-3, no growth factors; (2) SKOV-3, EGF 15'; (3) SKOV-3, EGF 30'; (4) SKOV-3, EGF 60'; (5) SKOV-3, Hrg 15'; (6) SKOV-3, Hrg 30'; (7) SKOV-3, Hrg 60'; (8) SKOV-3var, no growth factors; (9) SKOV-3var, serum 30'; (10) SKOV-3var, EGF 15'; (11) SKOV-3var, EGF 30'; (12) SKOV-3var, EGF 60'; (13) SKOV-3var, Hrg 15'; (14) SKOV-3var, Hrg 30'; (15) SKOV-3var, Hrg 60'; (16) HEY plastic; (17) HEY soft agar; (18) SKOV-3; (19) SKOV-3var; (20) A2780; (21) A2780-ADR; (22) OVCAR-3; (23) OVCAR-4; (24) MDA2774; (25) DOV13; (26) Caov-3; (27) ES-2; (28) HEY 0 hours; (29) HEY 1 hour; (30) HEY 3 hours; (31) HEY 6 hours; (32) HEY 9 hours; (33) HEY 12 hours; (34) SKOV-3 plastic; (35) SKOV-3 SubQ tumor; (36) SKOV-3 variant, plastic; (37) SKOV-3 var, SubQ tumor; (38) normal ovary; (39) normal ovary; (40) ovarian ascites; (41) ovarian ascites.

Table 1:

Sample Number	Tissue Type	Relative Expression
1	Artery normal	8.4607
2	Aorta diseased	10.4164
3	Vein normal	1.7542
4	Coronary SMC	0
5	HUVEC	341.5101
6	Heart normal	35.1581
7	Heart CHF	12.2591
8	Kidney	197.5103
9	Skeletal Muscle	24.6034
10	Adipose normal	4.6293
11	Pancreas	58.924
12	primary osteoblasts	0.269
13	Osteoclasts (diff)	1.8035
14	Skin normal	282.2411

15	Spinal cord normal	116.6291
16	Brain Cortex normal	2313.3764
17	Brain Hypothalamus normal	1500.039
18	Nerve	110.7209
19	DRG (Dorsal Root Ganglion)	119.908
20	Breast normal	9.7526
21	Breast tumor	6.7776
22	Ovary normal	26.278
23	Ovary Tumor	347.4796
24	Prostate Normal	147.6241
25	Prostate Tumor	144.0858
26	Salivary glands	10.0965
27	Colon normal	11.5577
28	Colon Tumor	62.5
29	Lung tumor	198.196
30	Lung COPD	22.2508
31	Colon IBD	58.517
32	Liver normal	30.9268
33	Liver fibrosis	8.5196
34	Spleen normal	1.3668
35	Tonsil normal	3.6195
36	Lymph node normal	4.1433
37	Small intestine normal	7.1146
38	Macrophages	1.3907
39	Synovium	0.7478
40	Activated PBMC	0.9698
41	Neutrophils	43.1349
42	Megakaryocytes	65.3803
43	Erythroid	181.1178
44	positive control	1337.9276

Table 2:

Sample Number	Tissue Type	Relative Expression
1	PIT 400 Breast N	3.17
2	PIT 372 Breast N	3.40
3	CHT 1228 Breast Normal	0.84
4	MDA 304 Breast T: MD-IDC	0.55
5	CHT 2002 Breast T: IDC	2.22

6	MDA 236-Breast T:PD-IDC(ILC?)	0.17
7	CHT 562 Breast T: IDC	0.52
8	NDR 138 Breast T ILC (LG)	0.36
9	CHT 1841 Lymph node (Breast met)	0.00
10	PIT 58 Lung (Breast met)	13.28
11	CHT 620 Ovary N	0.37
12	PIT 208 Ovary N	10.34
13	CLN 012 Ovary T	141.61
14	CLN 07 Ovary T	216.13
15	CLN 17 Ovary T	248.27
16	MDA 25 Ovary T	458.50
17	CLN 08 Ovary T	104.39
18	MDA 185 Lung N	7.14
19	PIT 298 Lung N	4.52
20	CLN 930 Lung N	8.82
21	MPI 215 Lung T--SmC	95.06
22	MDA 259 Lung T-PDNSCCL	88.08
23	CHT 832 Lung T-PDNSCCL	10.17
24	MDA 262 Lung T-SCC	18.58
25	CHT 793 Lung T-ACA	10.06
26	CHT 331 Lung T-ACA	15.68
27	CHT 405 Colon N	3.64
28	CHT 523 Colon N	3.42
29	CHT 371 Colon N	1.10
30	CHT 382 Colon T: MD	24.77
31	CHT 528 Colon T: MD	30.40
32	CLN 609 Colon T	1.41
33	NDR 210 Colon T: MD-PD	13.51
34	CHT 340 Colon-Liver Met	4.98
35	CHT 1637Colon-Liver Met	0.74
36	PIT 260 Liver N (female)	1.75
37	CHT 1653 Cervix Squamous CC	6.41
38	CHT 569 Cervix Squamous CC	0.25
39	A24 HMVEC-Arr	2.71
40	C48 HMVEC-Prol	2.85
41	Pooled Hemangiomas	0.12
42	HCT116N22 Normoxic	36.27

43	HCT116H22 Hypoxic	20.05

Table 3:

Sample Number	Tissue Type	Relative Expression
1	MCF-7 Breast T	25.56
2	ZR75 Breast T	14.48
3	T47D Breast T	11.40
4	MDA 231 Breast T	2.36
5	MDA 435 Breast T	2.65
6	SKBr3 Breast	10.86
7	DLD 1 ColonT (stageC)	57.31
8	SW480 Colon T (stage B)	2.88
9	SW620 ColonT (stageC)	13.23
10	HCT116	5.58
11	HT29	0.64
12	Colo 205	0.46
13	NCIH125	9.52
14	NCIH67	6.32
15	NCIH322	18.58
16	NCIH460	3.92
17	A549	10.10
18	NHBE	19.85
19	SKOV-3 ovary	4.06
20	OVCAR-3 ovary	11.84
21	293 Baby Kidney	27.30
22	293T Baby Kidney	44.66

Table 4:

Sample Number	Tissue Type	Relative Expression
1	MCF-7 Breast T	20.83
2	ZR75 Breast T	6.15
3	T47D Breast T	8.97
4	MDA 231 Breast T	0.29
5	MDA 435 Breast T	0.54
6	SKBr3 Breast	20.40
7	DLD 1 ColonT (stageC)	9.82
8	SW480 Colon T (stage B)	6.92

9	SW620 ColonT (stageC)	41.67
10	HCT116	26.01
11	HT29	3.04
12	Colo 205	1.44
13	NCIH125	12.05
14	NCIH67	458.50
15	NCIH322	6.11
16	NCIH460	3.27
17	A549	26.46
18	NHBE	0.00
19	SKOV-3 ovary	191.44
20	OVCAR-3 ovary	39.97
21	293 Baby Kidney	150.73
22	293T Baby Kidney	172.54

Table 5:

Sample Number	Tissue Type	Relative Expression
1	CHT 410 Colon N	1.59
2	CHT 425 Colon N	1.68
3	CHT 371 Colon N	0.16
4	PIT 281 Colon N	1.53
5	NDR 211 Colon N	0.40
6	CHT 122 Adenomas	3.33
7	CHT 887 Adenomas	10.20
8	CHT 414 Colonic ACA-B	2.13
9	CHT 841 Colonic ACA-B	0.39
10	CHT 890 Colonic ACA-B	1.67
11	CHT 910 Colonic ACA-B	6.78
12	CHT 377 Colonic ACA-B	0.48
13	CHT 520 Colonic ACA-C	1.89
14	CHT 596 Colonic ACA-C	2.09
15	CHT 907 Colonic ACA-C	5.08
16	CHT 372 Colonic ACA-C	7.57
17	NDR 210 Colonic ACA-C	0.21
18	CHT 1365 Colonic ACA-C	0.65
19	CLN 740 Liver N	0.19
20	CLN 741 Liver N	0.35
21	NDR 165 Liver N	0.81

22	NDR 150 Liver N	1.11
23	PIT 236 Liver N	1.04
24	CHT 1878 Liver N	1.30
25	CHT 119 Col Liver Met	13.32
26	CHT 131 Col Liver Met	0.95
27	CHT 218 Col Liver Met	1.21
28	CHT 739 Col Liver Met	0.84
29	CHT 755 Col Liver Met	0.61
30	CHT 215 Col Abdominal Met	2.13

Table 6:

Sample Number	Tissue Type	Relative Expression
1	CHT 410 Colon N	0.94
2	CHT 425 Colon N	0.80
3	CHT 371 Colon N	0.09
4	PIT 281 Colon N	0.79
5	NDR 211 Colon N	0.18
6	CHT 122 Adenomas	2.08
7	CHT 887 Adenomas	6.82
8	CHT 414 Colonic ACA-B	1.21
9	CHT 841 Colonic ACA-B	0.20
10	CHT 890 Colonic ACA-B	0.85
11	CHT 910 Colonic ACA-B	3.26
12	CHT 377 Colonic ACA-B	0.21
13	CHT 520 Colonic ACA-C	1.13
14	CHT 596 Colonic ACA-C	1.12
15	CHT 907 Colonic ACA-C	4.68
16	CHT 372 Colonic ACA-C	4.71
17	NDR 210 Colonic ACA-C	0.09
18	CHT 1365 Colonic ACA-C	0.32
19	CLN 740 Liver N	0.00
20	CLN 741 Liver N	0.17
21	NDR 165 Liver N	0.44
22	NDR 150 Liver N	0.47
23	PIT 236 Liver N	0.00
24	CHT 1878 Liver N	0.49
25	CHT 119 Col Liver Met	8.14

26	CHT 131 Col Liver Met	0.38
27	CHT 218 Col Liver Met	0.47
28	CHT 739 Col Liver Met	0.40
29	CHT 755 Col Liver Met	0.27
30	CHT 215 Col Abdominal Met	0.50

Table 7:

Sample Number	Tissue Type	Relative Expression
1	CHT 371 Colon N	0.02
2	CHT 523 Colon N	0.24
3	NDR 104 Colon N	1.88
4	CHT 520 Colonic ACA-C	0.77
5	CHT 1365 Colonic ACA-C	0.11
6	CHT 382 Colonic ACA-B	0.42
7	CHT 122 Adenocarcinoma	8.34
8	CHT 077 Liver-Colon Mets	2.42
9	CHT 739 Liver-Colon Mets	0.06
10	CHT 755 Liver-Colon Mets	0.08
11	CHT001 Liver-Colon Mets	0.23
12	CHT 084 Liver-Colon Mets	0.17
13	CHT 113 Liver-Colon Mets	0.11
14	CHT 114 Liver-Colon Mets	1.32
15	CHT 127 Liver-Colon Mets	0.87
16	CHT 137 Liver-Colon Mets	115.82
17	CHT 218 Liver-Colon Mets	0.16
18	CHT 220 Liver-Colon Mets	0.17
19	CHT 324 Liver-Colon Mets	1.28
20	CHT 340 Liver-Colon Met	6.78
21	CHT 530 Liver -Colon Met	0.18
22	CHT 849 Liver-Colon Met	7.02
23	CHT 1637 Liver-Colon Met	0.31
24	CHT131 Liver-Colon Met	0.38
25	NDR 165 Liver Normal	0.51
26	NDR 150 Liver Nomal	0.55
27	PIT 236 Liver Normal	0.46

Table 8:

Sample Number	Tissue Type	Relative Expression
1	JDLD-1	61.21
2	DKO1	31.14
3	DKS-8	21.20
4	DKO4	6.68
5	JHCT116	69.35
6	HK2-6	59.95
7	HKe3	25.38
8	HKh2	23.77
9	e3Ham#9	29.46

Table 9:

Sample Number	Tissue Type	Relative Expression
1	HCT 116 Aphidl t=0	29.16
2	HCT 116 Aphidl t=3	30.19
3	HCT 116 Aphidl t=6	27.39
4	HCT 116 Aphidl t=9	26.19
5	HCT 116 Aphidl t=12	24.86
6	HCT 116 Aphidl t=15	15.20
7	HCT 116 Aphidl t=18	25.74
8	HCT 116 Aphidl t=21	26.37
9	HCT 116 Aphidl t=24	19.51
10	HCT 116 Noc t=0	26.37
11	HCT 116 Noc t=3	20.91
12	HCT 116 Noc t=6	22.72
13	HCT 116 Noc t=9	26.74
14	HCT 116 Noc t=15	30.50
15	HCT 116 Noc t=18	44.19
16	HCT 116 Noc t=21	46.55
17	HCT 116 Noc t=24	41.67
18	DLD noc t=3	44.97
19	DLD noc t=6	62.07
20	DLD noc t=9	79.94
21	DLD noc t=12	73.81
22	DLD noc t=15	99.79
23	DLD noc t=18	70.56

24	DLD noc t=21	74.58
25	A549 Mimo t=0	67.45
26	A549 Mimo t=3	62.28
27	A549 Mimo t=6	45.12
28	A549 Mimo t=9	41.67
29	A549 Mimo t=15	33.73
30	A549 Mimo t=18	29.56
31	A549 Mimo t=21	28.26
32	A549 Mimo t=24	27.97
33	MCF10A Mimo t=0	0.00
34	MCF10A Mimo t=3	0.00
35	MCF10A Mimo t=6	0.00
36	MCF10A Mimo t=9	0.00
37	MCF10A Mimo t=12	0.00
38	MCF10A Mimo t=18	0.00
39	MCF10A Mimo t=21	0.00
40	MCF10A Mimo t=24	0.00

Table 10:

Sample Number	Tissue Type	Relative Expression
1	SMAD4-SW480 C	2.17
2	SMAD4-SW480 24HR	4.27
3	SMAD4-SW480 48HR	4.91
4	SMAD4-SW480 72HR	1.62
5	L51747-MUCINOUS	12.56
6	HT29 NON-MUCINOUS	3.66
7	SW620 NON-MUCINOUS	8.88
8	CSC-1 NORMAL	5.08
9	NCM-460 NORMAL	8.40
10	HCT116 RER+	14.89
11	SW48 RER+	20.69
12	SW480 RER-/-	9.82
13	CACO- RER-/-	4.89
14	JDLD-1	61.21
15	JHCT116	69.35
16	DKO1	31.14
17	DKO4	6.68
18	DKS-8	21.20

19	HKe3	25.38
20	HKh2	23.77
21	HK2-6	59.95
22	e3Ham#9	29.46
23	APC5 -/-	0.00
24	APC6-/-	2.00
25	APC1+/+	0.31
26	APC13+/+	0.61

Table 11:

Sample Number	Tissue Type	Relative Expression
1	SKOV-3 Serum '30	71.1
2	SKOV-3 No GF	86.9
3	SKOV-3 EGF '15	82.8
4	SKOV-3 EGF '30	75.9
5	SKOV-3 EGF '60	66.3
6	SKOV-3 Hrg '15	64.9
7	SKOV-3 Hrg '30	63.4
8	SKOV-3 Hrg '60	67.9
9	SKOV-3var No GF	76.9
10	SKOV-3var Serum '30	98.8
11	SKOV-3var EGF '15	58.7
12	SKOV-3var EGF '30	60.2
13	SKOV-3var EGF '60	46.9
14	SKOV-3var Hrg 15'	62.3
15	SKOV-3var Hrg 30'	96.4
16	SKOV-3var Hrg 60'	71.8
17	HEY Plastic	0.6
18	HEY Soft Agar	0.0
19	SKOV-3	66.3
20	SKOV-3var	32.5
21	A2780	32.9
22	A2780-ADR	7.9
23	OVCAR-3	40.1
24	OVCAR-4	6.9
25	MDA2774	37.9
26	DOV13	17.8
27	Caov-3	6.7

28	ES-2	0.0
29	HEY 0hr	2.6
30	HEY 1hr	2.5
31	HEY 3hr	3.6
32	HEY 6hr	2.8
33	HEY 9hr	2.6
34	HEY 12hr	1.4
35	SKOV-3 Plastic	167.2
36	SKOV-3 SubQ Tumor	19.1
37	SKOV-3 Variant Plastic	135.8
38	SKOV-3 Var SubQ Tumor	1.5
39	MDA 127 Normal Ovary	0.3
40	MDA 224 Normal Ovary	0.1
41	MDA 124 Ovarian Ascites	0.2
42	MDA 126 Ovarian Ascites	53.8

The present invention provides methods and compositions for the diagnosis and treatment of cellular proliferation disorders, *e.g.*, breast cancer, ovarian cancer, lung cancer, and/or colon cancer. The present invention is based, at least in part, on the discovery that expression of the 25943 gene (a glycosylasparaginase) is upregulated in tumors (*e.g.*, ovarian, lung, breast, and colon tumors). The present invention is further based, at least in part, on the discovery that 25943 expression is regulated during the cell cycle. The invention is still further based, at least in part, on the discovery that 25943 may be involved in the post-translational modification and processing of proteins. Without intending to be limited by mechanism, it is believed that modulation, *e.g.*, inhibition, of 25943 activity may modulate, *e.g.*, inhibit, protein modulation (*e.g.*, post-translational modification and process) in tumor cells that is relevant to their tumorigenic potential, therefore, modulating, *e.g.*, inhibiting cellular proliferation and tumorigenesis.

25943 is a member of a class of enzymes called glycosylasparaginases, which are involved in the degradation of glycoproteins, predominantly in the lysosomes. Generally, deficiencies in these types of enzymes cause the accumulation of partially degraded oligosaccharides and glycopeptides. Deficiencies in glycosylasparaginase, also referred to as aspartylglucosaminidase, result in the lysosomal storage disease known as Aspartylglucosaminuria. The glycosylasparaginase family of enzymes has multifunctional properties and wide substrate specificity. Glycosylasparaginases are responsible for catalyzing the cleavage of the protein-to-carbohydrate linkage of asparagine-linked

glycoproteins. Glycosylasparaginases can also catalyze the hydrolysis of β-aspartyl peptides to form aspartic acid and amino acids or peptides. Aspartate, a 4-carbon amino acid, is then transformed to oxaloacetate, which is used to begin the citric acid cycle, a major pathway for the generation of ATP and a provider of intermediates for the synthesis of amino acids. The general reaction that glycosylasparaginases catalyze is as follows:

N4-(β-N-acetyl-D-glucosaminyl)-L-asparagine + H_2O =
N-acetyl-β-glucosaminylamine + L-aspartate

The 25943 modulators identified according to the methods of the invention can be used to modulate cellular proliferation (*e.g.*, in breast, ovary, lung, and/or colon cells) and are, therefore, useful in treating, diagnosing, or prognosing cellular proliferation disorders. For example, inhibition of the activity of a 25943 molecule can inhibit cellular proliferation, thereby inhibiting tumorigenesis in the subject. Thus, the 25943 modulators identified using the assays described herein can be used to treat cellular proliferation disorders (*e.g.*, cancer) and/or disorders which are secondary to such disorders. Alternatively, 25943 modulators can increase cellular proliferation by increasing 25943 activity in a subject. Thus, 25943 modulators are also useful in the treatment of undesirable cell death, *e.g.*, neurodegenerative disorders.

As used herein, "cellular proliferation disorders" include those disorders that affect cellular proliferation, growth, apoptosis, differentiation, and/or migration processes. As used herein, a "cellular proliferation, growth, apoptosis, differentiation, and/or migration process" is a process by which a cell increases in number, size or content, by which a cell undergoes programmed cell death, by which a cell develops a specialized set of characteristics which differ from that of other cells, or by which a cell moves closer to or further from a particular location or stimulus. Examples of cellular proliferation disorders include cancer, *e.g.*, breast cancer, colon cancer, lung cancer, ovarian cancer, as well as other types of carcinomas, sarcomas, lymphomas, and/or leukemias; tumor angiogenesis and metastasis; skeletal dysplasia; hepatic disorders; and hematopoietic and/or myeloproliferative disorders. Other examples of disorders characterized by aberrant regulation of apoptosis include stroke-associated cell death and neurodegenerative disorders such as Alzheimer's disease, dementias related to Alzheimer's disease (such as

Pick's disease), Parkinson's and other Lewy diffuse body diseases, senile dementia, and Huntington's disease.

As used interchangeably herein, "25943 activity," "biological activity of 25943" or "functional activity of 25943," includes an activity exerted by a 25943 protein, polypeptide or nucleic acid molecule on a 25943 responsive cell or tissue (*e.g.*, breast, ovary, lung, or colon) or on a 25943 protein substrate, as determined *in vivo*, or *in vitro*, according to standard techniques. 25943 activity can be a direct activity, such as an association with a 25943-target molecule. As used herein, a "substrate" or "target molecule" or "binding partner" is a molecule with which a 25943 protein binds or interacts in nature, such that 25943-mediated function, *e.g.*, modulation of protein-carbohydrate linkage, is achieved. A 25943 target molecule can be a non-25943 molecule (*e.g.*, a carbohydrate-linked protein), or a 25943 protein or polypeptide. Examples of such target molecules include proteins in the same signaling path as the 25943 protein, *e.g.*, proteins which may function upstream (including both stimulators and inhibitors of activity) or downstream of the 25943 protein in a pathway involving regulation of cellular proliferation. Alternatively, a 25943 activity is an indirect activity, such as a cellular signaling activity mediated by interaction of the 25943 protein with a 25943 target molecule. The biological activities of 25943 are described herein. For example, the 25943 proteins can have one or more of the following activities: 1) they modulate hydrolysis of the protein-carbohydrate linkage of asparagine-linked glycoproteins; 2) they modulate hydrolysis of β-aspartyl peptides into aspartic acid and amino acids and/or peptides; 3) they modulate the citric acid cycle; 4) the modulate posttranslational modification and/or processing of proteins; and/or 5) they modulate cellular proliferation, growth, apoptosis, differentiation, and/or migration (*e.g.*, in breast, ovary, lung, and/or colon cells).

Various aspects of the invention are described in further detail in the following subsections:

I. Screening Assays:

The invention provides methods (also referred to herein as "screening assays") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules, ribozymes, or 25943 antisense molecules) which bind to 25943 proteins, have a stimulatory or inhibitory effect on 25943 expression or 25943 activity, or have a stimulatory or inhibitory effect on the expression or activity of a 25943

target molecule. Compounds identified using the assays described herein may be useful for treating cellular proliferation disorders.

Candidate/test compounds include, for example, 1) peptides such as soluble peptides, including Ig-tailed fusion peptides and members of random peptide libraries (see, *e.g.*, Lam, K.S. *et al.* (1991) *Nature* 354:82-84; Houghten, R. *et al.* (1991) *Nature* 354:84-86) and combinatorial chemistry-derived molecular libraries made of D- and/or L-configuration amino acids; 2) phosphopeptides (*e.g.*, members of random and partially degenerate, directed phosphopeptide libraries, see, *e.g.*, Songyang, Z. *et al.* (1993) *Cell* 72:767-778); 3) antibodies (*e.g.*, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric, and single chain antibodies as well as Fab, $F(ab')_2$, Fab expression library fragments, and epitope-binding fragments of antibodies); and 4) small organic and inorganic molecules (*e.g.*, molecules obtained from combinatorial and natural product libraries).

The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. (1997) *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994) *J. Med. Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and Gallop *et al.* (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP '409), plasmids (Cull *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:1865-1869) or phage (Scott and Smith (1990) *Science* 249:386-390; Devlin (1990) *Science* 249:404-406; Cwirla *et al.*

(1990) *Proc. Natl. Acad. Sci.* 87:6378-6382; Felici (1991) *J. Mol. Biol.* 222:301-310; Ladner *supra.*).

In one aspect, an assay is a cell-based assay in which a cell which expresses a 25943 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate 25943 activity is determined. In a preferred embodiment, the biologically active portion of the 25943 protein includes a domain or motif which can modulate hydrolysis of a protein-carbohydrate linkage. Determining the ability of the test compound to modulate 25943 activity can be accomplished by monitoring, for example, the production of one or more specific metabolites (*e.g.*, free carbohydrate, Asn-GlcNAc, aspartic acid, amino acids, and/or peptides), by measuring expression of cell cycle regulatory genes, or by monitoring cellular proliferation. The cell, for example, can be of mammalian origin, *e.g.*, a breast cell, a lung cell, an ovary cell, or a colon cell.

The ability of the test compound to modulate 25943 binding to a substrate can also be determined. Determining the ability of the test compound to modulate 25943 binding to a substrate (*e.g.*, an Asn-linked glycoprotein) can be accomplished, for example, by coupling the 25943 substrate with a radioisotope, fluorescent, or enzymatic label such that binding of the 25943 substrate to 25943 can be determined by detecting the labeled 25943 substrate in a complex. Alternatively, 25943 could be coupled with a radioisotope or enzymatic label to monitor the ability of a test compound to modulate 25943 binding to a 25943 substrate in a complex. Determining the ability of the test compound to bind 25943 can be accomplished, for example, by coupling the compound with a radioisotope or enzymatic label such that binding of the compound to 25943 can be determined by detecting the labeled 25943 compound in a complex. For example, 25943 substrates can be labeled with ^{125}I, ^{35}S, ^{14}C, or ^{3}H, either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product.

It is also within the scope of this invention to determine the ability of a compound to interact with 25943 without the labeling of any of the interactants. For example, a microphysiometer can be used to detect the interaction of a compound with 25943 without the labeling of either the compound or the 25943 (McConnell, H.M. *et al.* (1992) *Science*

257:1906-1912). As used herein, a "microphysiometer" (*e.g.*, Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a compound and 25943.

Because 25943 expression is increased in tumors, including metastatic tumors, and is regulated during the cell cycle, compounds which modulate cellular proliferation can be identified by the ability to modulate 25943 expression. To determine whether a test compound modulates 25943 expression, a cell which expresses 25943 (*e.g.*, a breast tumor cell, a lung tumor cell, an ovary tumor cell, a colon tumor cell, or a corresponding normal cell) is contacted with a test compound, and the ability of the test compound to modulate 25943 expression can be determined by measuring 25943 mRNA by, *e.g.*, Northern Blotting, quantitative PCR (*e.g.*, Taqman), or *in vitro* transcriptional assays. To perform an *in vitro* transcriptional assay, the full length promoter and enhancer of 25943 can be linked to a reporter gene such as chloramphenicol acetyltransferase (CAT) or luciferase and introduced into host cells. The same host cells can then be transfected with or contacted with the test compound. The effect of the test compound can be measured by reporter gene activity and comparing it to reporter gene activity in cells which do not contain the test compound. An increase or decrease in reporter gene activity indicates a modulation of 25943 expression and is, therefore, an indicator of the ability of the test compound to modulate cellular proliferation.

The ability of the test compound to modulate 25943 expression can also be determined by measuring the glycosylasparaginase activity present in a cell contacted with a test compound. To determine whether a test compound modulates 25943 glycosylasparaginase activity, a cell which expresses 25943 (*e.g.*, a breast tumor cell, a lung tumor cell, an ovary tumor cell, a colon tumor cell, or a corresponding normal cell) is contacted with a test compound, and the ability of the test compound to modulate 25943 glycosylasparaginase activity can be determined by measuring the level of GlcNAc, for example. Exemplary methods for measuring 25943 glycosylasparaginase activity are described in detail in Examples 4-7 and 9-10.

Cell lines transiently and stably transfected with tumor suppressors and oncogenes known to be associated with colon cancer progression may be useful in the methods of the invention for the identification of 25943 modulators (*e.g.*, SW480 cells stably or transiently transfected with Smad4). Smad4 is a candidate tumor suppressor gene mutated in a subset of colon carcinomas. Smad4 functions in the signal transduction of TGF-β

molecules. It is well known that the TGF-β superfamily is involved in growth inhibition. Smad4 mutation/loss in colon cell lines provides the hypothesis that Smad4 may be a modulator of cell adhesion and invasion. Other cell lines useful in the methods of the invention are NCM425 cells stably or transiently transfected with β-catenin. Mutations of the APC gene are responsible for tumor formation in sporadic and familial forms of colorectal cancer. APC binds β-catenin and regulates the cytoplasmic levels of β-catenin. When APC is mutated, β-catenin accumulates in the cytoplasm and translocates into the nucleus. Once in the nucleus it interacts with LEF/TCF molecules and regulates gene expression. Genes regulated by the β-catenin/LEF complex, like c-myc and cyclin D1, are involved in tumorigenesis. Also useful in the methods of the invention are cells stably or transiently transfected with p53. p53 is a well known tumor suppressor which is mutated in >50% of colorectal cancer tumors.

Abnormalities in cell cycle regulation and its checkpoints lead to the development of malignant cells. The loss of a cell's ability to respond to signals that regulate cell proliferation and cell cycle arrest is a common mechanism of cancer. Accordingly, for the study of specific time point within the cell cycle, cell lines such as the colon cancer cell lines HCT116, DLD-1 and NCM425 may be synchronized with agents such as Aphidicolin (G1 block), Mimosine (G1 block) and Nocodazole (G2/M block).

Other cell lines useful in the methods of the invention included the colon cancer cell lines HCT116 and DLD1 with disrupted k-ras genes. Point mutations that activate the k-ras oncogene are found in 50% of human colon cancers. Activated k-ras may be regulating cell proliferation in colorectal tumors. Disrupting the activated k-ras allele in HCT116 and DLD1 cells morphologically alters differentiation, causes loss of anchorage independent growth, slows proliferation *in vitro* and *in vivo*, and reduces expression of c-myc. Still other cell lines useful in the methods of the invention include transient or stable transfections of WISP-1 into NCM425 colon cancer cells, transient or stable transfections of DCC, Cox2, and/or APC into various cells.

Assays that may be used to identify compounds that modulate 25943 activity also include assays that test for the ability of a compound to modulate cellular proliferation. The ability of a test compound to modulate cellular proliferation can be measured by its ability to modulate proliferation in a cell which expresses 25943, *e.g.*, a breast, ovary, lung, or colon cell such as a breast, ovary, lung, or colon tumor cell. For example, the ability of a test compound to modulate cellular proliferation can be measured by contacting

a cell (*e.g.*, a breast, ovary, lung, or colon tumor cell) with the test compound, incubating the cell for a period of time, and measuring the number of cells present as compared to a control cell not contacted with the test compound. The number of cells can be measured, for example, by dry/wet weight measurement (see Example 1), by counting the cells via optical density (see Example 2), by using a counting chamber (see Example 3), or by using a Coulter Counter. The ability of a test compound to modulate cellular proliferation can also be measured by contacting a cell (*e.g.*, a breast, ovary, lung, or colon tumor cell) with the test compound and testing the ability of the cell to form a colony in soft agar (see Example 8). The ability of a cell to grow in soft agar indicates that it has lost the requirement for anchorage-dependant growth, which is an indication of tumorigenic potential. The ability of a test compound to modulate cellular proliferation may also be measured by contacting a cell (*e.g.*, a breast, ovary, lung, or colon tumor cell) with the test compound and testing the ability of the cell to form a tumor in a nude mouse. The nude mouse, a hairless mutant discovered in 1962, is immunodeficient, and thus does not reject tumor transplantations from other species.

Numerous other methods exist in the art to measure cellular proliferation. Examples include measurement of the metabolic activity of viable cells via WST-8 reduction to formazan salt using a colorimetric assay (Cell Counting Kit-8 from Alexis Biochemicals, San Diego, CA or from Dojindo Molecular Technologies, Inc., Gaithersburg, MD); measurement of DNA synthesis by BrdU incorporation using an anti-BrdU monoclonal antibody/horseradish peroxidase-based detection system (Cell Proliferation ELISA or Immunocytochemistry from Amersham Pharmacia Biotech, Piscataway, NJ); DNA synthesis by [^{14}C]thymidine uptake (Thymidine Uptake [^{14}C] Cytostar-T Assay from Amersham Pharmacia Biotech, Piscataway, NJ); and DNA synthesis measured by scintillation proximity assay (SPA) of [^{3}H]thymidine incorporation ([^{3}H]Thymidine Uptake Assay Kit from Amersham Pharmacia Biotech, Piscataway, NJ).

Further examples of methods for measuring cellular proliferation include measurement of simultaneous cell surface markers and intracellular BrdU incorporation (FastImmune Anti-BrdU with DNase from BD Biosciences, San Jose, CA); measurement of the metabolic activity of viable cells via WST-1 reduction to soluble formazan salt using a colorimetric assay (Quick Cell Proliferation Assay Kit from BioVision, Inc., Mountain View, CA; Cell Proliferation Assay Kit from Chemicon International, Inc., Temecula, CA; Rapid Cell Viability Assay from Oncogene Research Products, San Diego, CA; Cell Proliferation Reagent WST-1 from R&D Systems, Minneapolis, MN);

measurement of live cells stained with "Cyto-dye" and dead cells stained with propidium iodide (Live/Dead Cell Staining Kit from BioVision, Inc., Mountain View, CA); and measurement of metabolic activity using bioluminescent detection of ATP (ApoSENSOR ATP Determination Kit from BioVision, Inc., Mountain View, CA; LumiTech's ViaLight HS Assay, LumiTech's ViaLight HT Assay, and LumiTech's ViaLight MDA Assay, all from BioWhittaker, Walkersville, MD; CytoLux Assay Kit from Perkin Elmer Life Sciences, Boston, MA; Cytotoxicity and Cell Proliferation Kit from Thermo Labsystems, Franklin, MA).

Additional examples of methods for measuring cellular proliferation include measurement of metabolic activity of viable cells via MTT reduction to formazan salt using a colorimetric assay (MTT Cell Growth Assay Kit from Chemicon International, Inc., Temecula, CA; Vybrant MTT Cell Proliferation Assay Kit from Molecular Probes, Inc., Eugene, OR; CellTiter 96 Non-Radioactive Cell Proliferation Assay from Promega, Madison, WI; TACS MTT Cell Proliferation and Viability Assay and Cell Proliferation Kit I MTT, both from R&D Systems, Minneapolis, MN; In Vitro Toxicology Assay Kit, MTT based from Sigma-Aldrich, St. Louis, MO); measurement of live cells stained with calcein-AM and dead cells labeled with propidium iodide (Cellstain Double-Staining Kit from Dojindo Molecular Technologies, Inc., Gaithersburg, MD); measurement of DNA content using CyQUANT GR dye (CyQUANT Cell Proliferation Assay Kit from Molecular Probes, Inc., Eugene, OR); measurement of DNA synthesis by BrdU incorporation using ELISA-based chemiluminescent detection (BrdU Cell Proliferation Assay from Oncogene Research Products, San Diego, CA; Cell Proliferation ELISA, BrdU (chemiluminescent) from R&D Systems, Minneapolis, MN); and measurement of DNA synthesis by BrdU incorporation using ELISA-based colorimetric detection (BrdU Proliferation Assay - HTS from Oncogene Research Products, San Diego, CA; BrdU Labeling and Detection Kit III and Cell Proliferation ELISA, BrdU (colorimetric), both from R&D Systems, Minneapolis, MN).

Further examples of methods for measuring cellular proliferation include measurement of proliferating cell nuclear antigen (PCNA) using biotinylated anti-PCNA monoclonal antibody (PCNA (Proliferating Cell Nuclear Antigen) ELISA from Oncogene Research Products, San Diego, CA); measurement of DNA synthesis by BrdU incorporation detection using an anti-BrdU monoclonal antibody (BrdU IHC System from Oncogene Research Products, San Diego, CA; BrdU Kit from Zymed Laboratories, Inc., South San Francisco, CA); measurement of DNA synthesis by BrdU incorporation using

Strand Break Induced Photolysis (SBIP) methodology, with break sites identified by BrdU incorporation (ABSOLUTE-S SBIP Cell Proliferation Assay Kit from Phoenix Flow Systems Inc., San Diego, CA); measurement of metabolic activity of viable cells via MTS reduction to soluble formazan salt using a colorimetric assay (CellTiter 96 Aqueous Non-Radioactive Cell Proliferation Assay from Promega, Madison, WI); and measurement of metabolic activity of viable cells via MTS reduction to formazan salt using a colorimetric assay (CellTiter 96 Aqueous One Solution Cell Proliferation Assay from Promega, Madison, WI).

Additional examples of methods for measuring cellular proliferation include measurement of metabolic activity via bioluminescent of ATP using luciferin and thermostable luciferase (CellTiter-Glo Luminescent Cell Viability Assay from Promega, Madison, WI); measurement of single-cell proliferation by direct immunofluorescence staining (In Situ Cell Proliferation Kit, FLUOS, and BrdU Labeling and Detection Kit I, both from R&D Systems, Minneapolis, MN) or indirect immunostaining method (BrdU Labeling and Detection Kit II from R&D Systems, Minneapolis, MN); measurement of metabolic activity of viable cells via XTT reduction to soluble formazan salt using a colorimetric assay (R&D Systems, Minneapolis, MN; In Vitro Toxicology Assay Kit, XTT based from Sigma-Aldrich, St. Louis, MO); detection of nuclear cell cycle-associated antigens expressed only in proliferating cells (Monoclonal Antibodies to Cell Cycle-Associated Antigens from R&D Systems, Minneapolis, MN); measurement of cell proliferation using plasma membrane dye (Cell Census Plus System from Sigma-Aldrich, St. Louis, MO); and measurement of membrane-associated phosphatase activity via conversion of p-nitrophenyl phosphate to a colored compound (In Vitro Toxicology Assay Kit, Acid Phosphatase based from Sigma-Aldrich, St. Louis, MO).

Still further examples of methods for measuring cellular proliferation include measurement of neutral red dye staining of viable cells using a colorimetric assay (In Vitro Toxicology Assay Kit, Neutral Red based from Sigma-Aldrich, St. Louis, MO); measurement of total protein upon sulforhodamine dye binding using a colorimetric assay (In Vitro Toxicology Assay Kit, Sulforhodamine B based from Sigma-Aldrich, St. Louis, MO); measurement of metabolic activity of viable cells measured by tetrazolium reduction to formazan derivative using a colorimetric assay (In Vitro Toxicology Assay Kit, Lactic Dehydrogenase based from Sigma-Aldrich, St. Louis, MO); measurement of metabolic activity of viable cells via the bioreduction of dye that converts the oxidized form (blue) to a fluorescent intermediate (red) (In Vitro Toxicology Assay Kit, Resazurin based from

Sigma-Aldrich, St. Louis, MO); measurement of DNA content using Quantos dye reagent (Quantos Cell Proliferation Assay Kit from Stratagene, La Jolla, CA); and measurement of DNA content by A:T base pair-binding dye (TACS Hoechst Cell Proliferation Assay I (CPA1) and TACS Hoechst Cell Proliferation Assay 2 (CPA2), both from Trevigen, Inc., Gaithersburg, MD).

In yet another embodiment, an assay of the present invention is a cell-free assay in which a 25943 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to or to modulate (*e.g.*, stimulate or inhibit) the activity of the 25943 protein or biologically active portion thereof is determined. Preferred biologically active portions of the 25943 proteins to be used in assays of the present invention include fragments which participate in interactions with non-25943 molecules, *e.g.*, fragments with high surface probability scores. Binding of the test compound to the 25943 protein can be determined either directly or indirectly as described above. Determining the ability of the 25943 protein to bind to a test compound can also be accomplished using a technology such as real-time Biomolecular Interaction Analysis (BIA) (Sjolander, S. and Urbaniczky, C. (1991) *Anal. Chem.* 63:2338-2345; Szabo *et al.* (1995) *Curr. Opin. Struct. Biol.* 5:699-705). As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the optical phenomenon of surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In yet another embodiment, the cell-free assay involves contacting a 25943 protein or biologically active portion thereof with a known compound which binds the 25943 protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the 25943 protein, wherein determining the ability of the test compound to interact with the 25943 protein comprises determining the ability of the 25943 protein to preferentially bind to or modulate the activity of a 25943 target molecule (*e.g.*, a 25943 substrate).

The cell-free assays of the present invention are amenable to use of both soluble and/or membrane-bound forms of isolated proteins (*e.g.*, 25943 proteins or biologically active portions thereof). In the case of cell-free assays in which a membrane-bound form of an isolated protein is used it may be desirable to utilize a solubilizing agent such that the membrane-bound form of the isolated protein is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-

dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)$_n$, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl=N,N-dimethyl-3-ammonio-1-propane sulfonate.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either 25943 or a 25943 target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to a 25943 protein, or interaction of a 25943 protein with a 25943 target molecule in the presence and absence of a test compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/25943 fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or 25943 protein, and the mixture incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components, the matrix is immobilized in the case of beads, and complex formation is determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of 25943 binding or activity determined using standard techniques.

Other techniques for immobilizing proteins or cell membrane preparations on matrices can also be used in the screening assays of the invention. For example, either a 25943 protein or a 25943 target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated 25943 protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies which are reactive with 25943 protein or target molecules but which do not interfere with binding of the 25943 protein to its target molecule can be derivatized to the wells of the plate, and

unbound target or 25943 protein is trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the 25943 protein or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the 25943 protein or target molecule.

In yet another aspect of the invention, the 25943 protein or fragments thereof can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300) to identify other proteins which bind to or interact with 25943 ("25943-binding proteins" or "25943-bp) and are involved in 25943 activity. Such 25943-binding proteins are also likely to be involved in the propagation of signals by the 25943 proteins or 25943 targets as, for example, downstream elements of a 25943-mediated signaling pathway. Alternatively, such 25943-binding proteins are likely to be 25943 inhibitors.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a 25943 protein is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo,* forming a 25943-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the 25943 protein.

In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell-free assay, and the ability of the agent to modulate the activity of a 25943 protein can be confirmed *in vivo*, *e.g.*, in an animal such as an animal model for tumorigenesis, as described elsewhere herein. Additionally, animals deficient in 25943

(*e.g.*, 25943 knockout mice) may be deficient in the ability to modulate cellular proliferation via a 25943-regulated pathway, and therefore may be useful in determining whether a test compound can modulate proliferation by bypassing 25943 and directly modulating the activity of downstream components of the pathway.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, an 25943 modulating agent, an antisense 25943 nucleic acid molecule, an 25943-specific antibody, or an 25943-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

For example, the ability of the agent to modulate the activity of an 25943 protein can be tested in an animal such as an animal model for a cellular proliferation disorder, *e.g.*, tumorigenesis. Animal based models for studying tumorigenesis *in vivo* are well known in the art (reviewed in Animal Models of Cancer Predisposition Syndromes, Hiai, H. and Hino, O. (eds.) 1999, *Progress in Experimental Tumor Research*, Vol. 35; Clarke, A.R. (2000) *Carcinogenesis* 21:435-41) and include, for example, carcinogen-induced tumors (Rithidech, K. *et al.* (1999) *Mutat. Res.* 428:33-39; Miller, M.L. *et al.* (2000) *Environ. Mol. Mutagen.* 35:319-327), injection and/or transplantation of tumor cells into an animal, as well as animals bearing mutations in growth regulatory genes, for example, oncogenes (*e.g.*, ras) (Arbeit, J.M. *et al.* (1993) *Am. J. Pathol.* 142:1187-1197; Sinn, E. *et al.* (1987) *Cell* 49:465-475; Thorgeirsson, SS *et al.* (2000) *Toxicol. Lett.* 112-113:553-555) and tumor suppressor genes (*e.g.*, p53) (Vooijs, M. *et al.* (1999) *Oncogene* 18:5293-5303; Clark A.R. (1995) *Cancer Metast. Rev.* 14:125-148; Kumar, T.R. *et al.* (1995) *J. Intern. Med.* 238:233-238; Donehower, L.A. *et al.* (1992) *Nature* 356215-221). Furthermore, experimental model systems are available for the study of, for example, ovarian cancer (Hamilton, T.C. *et al.* (1984) *Semin. Oncol.* 11:285-298; Rahman, N.A. *et al.* (1998) *Mol. Cell. Endocrinol.* 145:167-174; Beamer, W.G. *et al.* (1998) *Toxicol. Pathol.* 26:704-710), gastric cancer (Thompson, J. *et al.* (2000) *Int. J. Cancer* 86:863-869; Fodde, R. *et al.* (1999) *Cytogenet. Cell Genet.* 86:105-111), breast cancer (Li, M. *et al.* (2000) *Oncogene* 19:1010-1019; Green, J.E. *et al.* (2000) *Oncogene* 19:1020-1027), melanoma

(Satyamoorthy, K. *et al.* (1999) *Cancer Metast. Rev.* 18:401-405); lung cancer (Malkinson, A.M. (2001) *Lung Cancer* 32(3):265-79; Zhao, B. *et al.* (2001) *Exp. Lung Res.* 26(8):567-79); colon cancer (Taketo, M.M. and Takaku (2000) *Hum. Cell* 13(3):85-95; Fodde, R. and Smits, R. (2001) *Trends. Mol. Med.* 7(8):369-73); and prostate cancer (Shirai, T. *et al.* (2000) *Mutat. Res.* 462:219-226; Bostwick, D.G. *et al.* (2000) *Prostate* 43:286-294). Other animal models which may be useful in the methods of the invention include mice with null mutations in the glycosylasparaginase gene (Kaartinen, V. *et al.* (1996) *Nat. Med.* 2:1375-1378).

Additional examples of mouse models for cancer are detailed below. For example, the Apc^{min} mouse is the most thoroughly characterized genetic model of human colorectal carcinogenesis. This model provides a valuable tool for identifying changes in gene expression associated with early stage disease resulting from the loss of Apc gatekeeper function. Adenomatous polyps and normal colonic epithelium from these mice may be harvested for standard and subtracted cDNA library construction and probe generation for microarray analysis. The Apc^{1638N} mouse was generated by introducing a PGK-neomycin gene at codon 1638 of the Apc gene. After 6-8 weeks, these mice form aberrant crypt foci which ultimately progress to carcinomas by 4 months of age. These mice on average develop 5-6 tumors within the upper gastrointestinal tract. In addition, these mice also develop extraintestinal tumors and desmoids. This lineage provides a means of studying extracolonic manifestations seen in familial adenomatous polyposis (FAP) patients such as desmoid disease. The $Smad3^{-/-}$ mouse has recently been described as a useful and unique model for human colorectal carcinogenesis. $Smad3^{-/-}$ mice develop colon carcinomas that histopathologically resemble human disease. One advantage of this model is that samples from several stages of disease progression can be isolated, including normal epithelium, hyperplastic epithelium, adenomatous polyps, and various degrees of primary carcinoma and lymph node metastases. Thus, the generation of subtracted cDNA libraries and probes representing these stages are a powerful tool for identifying and validating colon cancer targets.

Also useful in the methods of the invention are mis-match repair models (MMRs). Hereditary nonpolyposis colon cancer (HNPCC), which is caused by germline mutations in MSH2 & MLH1, genes involved in DNA mismatch repair, accounts for 5-15% of colon cancer cases. Mouse models have been generated carrying null mutations in the MLH1, MSH2 and MSH3 genes.

Xenograft mouse models are made by grafting cells from colon tumor cell lines into mice, *e.g.*, nude mice. Such genes could be crucial targets for anti-cancer drug development. Examples of colon tumor cell lines which may be used in the methods of the invention to create xenograft mouse models include HCT116, HT29, SW480, SW620, Colon 26, DLD1, Caco2, colo205, T84, CC-ML3, KM12C, KM12SM, HCC-2998, HCT-15, KM20L2, and KM12.

Examples of ovary tumor cell lines which may be used in the methods of the invention include cell lines SKOV3, SKOV3/Variant, OVCAR-3, OVCAR-4, and HEY. The SKOV3/Var cell line is a variant of the parental cell line SKOV3 that is resistant to cisplatin.

The HCT-116 human colon carcinoma cell line can be grown as a subcutaneous or orthotopic xenograft (intracaecal injection) in athymic nude mice, but metastasizes with low frequency. Rare liver and lung metastases can be isolated, expanded *in vitro*, and reimplanted *in vivo*. A limited (1-3) number of iterations of this process can be employed to isolate highly metastatic variants of the parental cell line. Standard and subtracted cDNA libraries and probes can be generated from the parental and variant cell lines to identify genes associated with the acquisition of a metastatic phenotype. This model can be established using several alternative human colon carcinoma cell lines, including SW480 and KM12C.

Additional animal models which may be useful in the methods of the invention are described in the Examples section herein.

In another aspect, cell-based systems, as described herein, may be used to identify compounds which may act to ameliorate tumorigenic or apoptotic disease symptoms. For example, such cell systems may be exposed to a compound, suspected of exhibiting an ability to ameliorate tumorigenic or apoptotic disease symptoms, at a sufficient concentration and for a time sufficient to elicit such an amelioration of tumorigenic or apoptotic disease symptoms in the exposed cells. After exposure, the cells are examined to determine whether one or more of the tumorigenic or apoptotic disease cellular phenotypes has been altered to resemble a more normal or more wild type, non-tumorigenic disease or non-apoptotic disease phenotype. Cellular phenotypes that are associated with tumorigenic disease states include aberrant proliferation and migration, angiogenesis, anchorage independent growth, and loss of contact inhibition. Cellular phenotypes that are associated with apoptotic disease states include aberrant DNA fragmentation, membrane blebbing, caspase activity, and cytochrome c release from mitochondria.

In addition, animal-based tumorigenic disease systems, such as those described herein, may be used to identify compounds capable of ameliorating tumorigenic or apoptotic disease symptoms. Such animal models may be used as test substrates for the identification of drugs, pharmaceuticals, therapies, and interventions which may be effective in treating tumorigenic or apoptotic disease. For example, animal models may be exposed to a compound, suspected of exhibiting an ability to ameliorate tumorigenic or apoptotic disease symptoms, at a sufficient concentration and for a time sufficient to elicit such an amelioration of tumorigenic or apoptotic tumorigenic or apoptotic disease symptoms in the exposed animals. The response of the animals to the exposure may be monitored by assessing the reversal of disorders associated with tumorigenic disease, for example, by counting the number of tumors and/or measuring their size before and after treatment. In addition, the animals may be monitored by assessing the reversal of disorders associated with tumorigenic disease, for example, reduction in tumor burden, tumor size, and invasive and/or metastatic potential before and after treatment.

With regard to intervention, any treatments which reverse any aspect of tumorigenic or apoptotic disease symptoms should be considered as candidates for human tumorigenic or apoptotic disease therapeutic intervention. Dosages of test agents may be determined by deriving dose-response curves.

Additionally, gene expression patterns may be utilized to assess the ability of a compound to ameliorate cardiovascular or tumorigenic disease symptoms. For example, the expression pattern of one or more genes may form part of a "gene expression profile" or "transcriptional profile" which may be then be used in such an assessment. "Gene expression profile" or "transcriptional profile", as used herein, includes the pattern of mRNA expression obtained for a given tissue or cell type under a given set of conditions. Such conditions may include, but are not limited to, the presence of a tumor, *e.g.*, a breast, colon, ovary, or lung tumor, including any of the control or experimental conditions described herein, for example, synchronized cells induced to enter the cell cycle, or RER- or Smad4 models. Other conditions may include, for example, cell differentiation, transformation, metastasis, and carcinogen exposure. Gene expression profiles may be generated, for example, by utilizing a differential display procedure, Northern analysis and/or RT-PCR. In one embodiment, 25943 gene sequences may be used as probes and/or PCR primers for the generation and corroboration of such gene expression profiles.

Gene expression profiles may be characterized for known states, either tumorigenic or apoptotic disease or normal, within the cell- and/or animal-based model systems.

Subsequently, these known gene expression profiles may be compared to ascertain the effect a test compound has to modify such gene expression profiles, and to cause the profile to more closely resemble that of a more desirable profile.

For example, administration of a compound may cause the gene expression profile of a tumorigenic or apoptotic disease model system to more closely resemble the control system. Administration of a compound may, alternatively, cause the gene expression profile of a control system to begin to mimic a tumorigenic or apoptotic disease state. Such a compound may, for example, be used in further characterizing the compound of interest, or may be used in the generation of additional animal models.

II. Predictive Medicine:

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining 25943 protein and/or nucleic acid expression as well as 25943 activity, in the context of a biological sample (*e.g.*, blood, serum, ascites, cells, or tissue, *e.g.*, breast, lung, colon, or ovarian tissue) to thereby determine whether an individual is afflicted with a cellular proliferation disorder. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a cellular proliferation disorder. For example, mutations in a 25943 gene can be assayed for in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a cellular proliferation disorder.

Another aspect of the invention pertains to monitoring the influence of 25943 modulators (*e.g.*, anti-25943 antibodies or 25943 ribozymes) on the expression or activity of 25943 in clinical trials.

These and other agents are described in further detail in the following sections.

A. Diagnostic Assays For Cellular Proliferation Disorders

To determine whether a subject is afflicted with a cellular proliferation disorder, a biological sample may be obtained from a subject and the biological sample may be contacted with a compound or an agent capable of detecting a 25943 protein or nucleic acid (*e.g.*, mRNA or genomic DNA) that encodes a 25943 protein, in the biological sample. A preferred agent for detecting 25943 mRNA or genomic DNA is a labeled

nucleic acid probe capable of hybridizing to 25943 mRNA or genomic DNA. The nucleic acid probe can be, for example, the 25943 nucleic acid set forth in SEQ ID NO:1, or a portion thereof, such as an oligonucleotide of at least 15, 20, 25, 30, 25, 40, 45, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to 25943 mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

A preferred agent for detecting 25943 protein in a sample is an antibody capable of binding to 25943 protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or F(ab')2) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of direct substances that can be coupled to an antibody or a nucleic acid probe include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin.

The term "biological sample" is intended to include tissues, cells, and biological fluids isolated from a subject, as well as tissues, cells, and fluids present within a subject. That is, the detection method of the invention can be used to detect 25943 mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of 25943 mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of 25943 protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of 25943 genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of 25943 protein include introducing into a subject a labeled anti-25943 antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting 25943 protein, mRNA, or genomic DNA, such that the presence of

25943 protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of 25943 protein, mRNA or genomic DNA in the control sample with the presence of 25943 protein, mRNA or genomic DNA in the test sample.

B. Prognostic Assays For Cellular Proliferation Disorder

The present invention further pertains to methods for identifying subjects having or at risk of developing a cellular proliferation disorder with aberrant 25943 expression or activity.

As used herein, the term "aberrant" includes a 25943 expression or activity which deviates from the wild type 25943 expression or activity. Aberrant expression or activity includes increased or decreased expression or activity, as well as expression or activity which does not follow the wild type developmental pattern of expression or the subcellular pattern of expression. For example, aberrant 25943 expression or activity is intended to include the cases in which a mutation in the 25943 gene causes the 25943 gene to be under-expressed or over-expressed and situations in which such mutations result in a non-functional 25943 protein or a protein which does not function in a wild-type fashion, *e.g.*, a protein which does not interact with a 25943 substrate, or one which interacts with a non-25943 substrate.

The assays described herein, such as the preceding diagnostic assays or the following assays, can be used to identify a subject having or at risk of developing a cellular proliferation disorder, *e.g.*, breast cancer, colon cancer, lung, cancer, and/or ovarian cancer. A biological sample may be obtained from a subject and tested for the presence or absence of a genetic alteration. For example, such genetic alterations can be detected by ascertaining the existence of at least one of 1) a deletion of one or more nucleotides from a 25943 gene, 2) an addition of one or more nucleotides to a 25943 gene, 3) a substitution of one or more nucleotides of a 25943 gene, 4) a chromosomal rearrangement of a 25943 gene, 5) an alteration in the level of a messenger RNA transcript of a 25943 gene, 6) aberrant modification of a 25943 gene, such as of the methylation pattern of the genomic DNA, 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a 25943 gene, 8) a non-wild type level of a 25943-protein, 9) allelic loss of a 25943 gene, and 10) inappropriate post-translational modification of a 25943-protein.

As described herein, there are a large number of assays known in the art which can be used for detecting genetic alterations in a 25943 gene. For example, a genetic alteration in a 25943 gene may be detected using a probe/primer in a polymerase chain reaction

(PCR) (see, *e.g.*, U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, *e.g.*, Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:360-364), the latter of which can be particularly useful for detecting point mutations in a 25943 gene (see Abravaya *et al.* (1995) *Nucleic Acids Res.* 23:675-682). This method includes collecting a biological sample from a subject, isolating nucleic acid (*e.g.*, genomic DNA, mRNA or both) from the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a 25943 gene under conditions such that hybridization and amplification of the 25943 gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli, J.C. *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh, D.Y. *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi, P.M. *et al.* (1988) *Bio-Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a 25943 gene from a biological sample can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in 25943 can be identified by hybridizing biological sample derived and control nucleic acids, *e.g.*, DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotide probes (Cronin, M.T. *et al.* (1996) *Hum. Mutat.* 7:244-255; Kozal, M.J. *et al.* (1996) *Nat. Med.* 2:753-759). For example, genetic mutations in 25943 can be identified in two dimensional arrays

containing light-generated DNA probes as described in Cronin, M.T. *et al.* (1996) *supra.* Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential, overlapping probes. This step allows for the identification of point mutations. This step is followed by a second hybridization array that allows for the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the 25943 gene in a biological sample and detect mutations by comparing the sequence of the 25943 in the biological sample with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxam and Gilbert (1977) *Proc. Natl. Acad. Sci. USA* 74:560) or Sanger (1977) *Proc. Natl. Acad. Sci. USA* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays (Naeve, C.W. (1995) *Biotechniques* 19:448-53), including sequencing by mass spectrometry (see, *e.g.*, PCT International Publication No. WO 94/16101; Cohen *et al.* (1996) *Adv. Chromatogr.* 36:127-162; and Griffin *et al.* (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

Other methods for detecting mutations in the 25943 gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes formed by hybridizing (labeled) RNA or DNA containing the wild-type 25943 sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digest the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for

example, Cotton *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:4397 and Saleeba *et al.* (1992) *Methods Enzymol.* 217:286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in 25943 cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a 25943 sequence, *e.g.*, a wild-type 25943 sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in 25943 genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl. Acad. Sci USA* 86:2766; see also Cotton (1993) *Mutat. Res.* 285:125-144 and Hayashi (1992) *Genet. Anal. Tech. Appl.* 9:73-79). Single-stranded DNA fragments of sample and control 25943 nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen *et al.* (1991) *Trends Genet.* 7:5).

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers *et al.* (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to ensure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in

place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys. Chem.* 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki *et al.* (1986) *Nature* 324:163); Saiki *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini *et al.* (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered a 25943 modulator (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, or small molecule) to effectively treat a cellular proliferation disorder.

C. Monitoring of Effects During Clinical Trials

The present invention further provides methods for determining the effectiveness of a 25943 modulator (*e.g.*, a 25943 modulator identified herein) in treating a cellular proliferation disorder in a subject. For example, the effectiveness of a 25943 modulator in

increasing 25943 gene expression, protein levels, or in upregulating 25943 activity, can be monitored in clinical trials of subjects exhibiting decreased 25943 gene expression, protein levels, or downregulated 25943 activity. Alternatively, the effectiveness of a 25943 modulator in decreasing 25943 gene expression, protein levels, or in downregulating 25943 activity, can be monitored in clinical trials of subjects exhibiting increased 25943 gene expression, protein levels, or 25943 activity. In such clinical trials, the expression or activity of a 25943 gene, and preferably, other genes that have been implicated in, for example, a cellular proliferation disorder can be used as a "read out" or marker of the phenotype of a particular cell.

For example, and not by way of limitation, genes, including 25943, that are modulated in cells by treatment with an agent which modulates 25943 activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents which modulate 25943 activity on subjects suffering from a cellular proliferation disorder in, for example, a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of 25943 and other genes implicated in the cellular proliferation disorder. The levels of gene expression (*e.g.*, a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods described herein, or by measuring the levels of activity of 25943 or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent which modulates 25943 activity. This response state may be determined before, and at various points during treatment of the individual with the agent which modulates 25943 activity.

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent which modulates 25943 activity (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, or small molecule identified by the screening assays described herein) including the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a 25943 protein, mRNA, or genomic DNA in the pre-administration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the 25943 protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the 25943 protein, mRNA, or genomic DNA in the pre-administration sample with the 25943 protein, mRNA, or genomic DNA in the post administration sample or

samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of 25943 to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of 25943 to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent. According to such an embodiment, 25943 expression or activity may be used as an indicator of the effectiveness of an agent, even in the absence of an observable phenotypic response.

III. Methods of Treatment of Subjects Suffering From Cellular Proliferation Disorders:

The present invention provides for both prophylactic and therapeutic methods of treating a subject, *e.g.*, a human, at risk of (or susceptible to) a cellular proliferation disorder such as breast cancer, ovarian cancer, lung cancer, and/or colon cancer. As used herein, "treatment" of a subject includes the application or administration of a therapeutic agent to a subject, or application or administration of a therapeutic agent to a cell or tissue from a subject, who has a diseases or disorder, has a symptom of a disease or disorder, or is at risk of (or susceptible to) a disease or disorder, with the purpose of curing, healing, alleviating, relieving, altering, remedying, ameliorating, improving, or affecting the disease or disorder, the symptom of the disease or disorder, or the risk of (or susceptibility to) the disease or disorder. As used herein, a "therapeutic agent" includes, but is not limited to, small molecules, peptides, polypeptides, antibodies, ribozymes, and antisense oligonucleotides.

With regard to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from the field of pharmacogenomics. "Pharmacogenomics," as used herein, refers to the application of genomics technologies such as gene sequencing, statistical genetics, and gene expression analysis to drugs in clinical development and on the market. More specifically, the term refers to the study of how a patient's genes determine his or her response to a drug (*e.g.*, a patient's "drug response phenotype", or "drug response genotype").

Thus, another aspect of the invention provides methods for tailoring a subject's prophylactic or therapeutic treatment with either the 25943 molecules of the present invention or 25943 modulators according to that individual's drug response genotype. Pharmacogenomics allows a clinician or physician to target prophylactic or therapeutic

treatments to patients who will most benefit from the treatment and to avoid treatment of patients who will experience toxic drug-related side effects.

A. Prophylactic Methods

In one aspect, the invention provides a method for preventing in a subject, a cellular proliferation disorder by administering to the subject an agent which modulates 25943 expression or 25943 activity, *e.g.*, modulation of cellular proliferation in, *e.g.*, breast, lung, colon, or ovary cells. Subjects at risk for a cellular proliferation disorder can be identified by, for example, any or a combination of the diagnostic or prognostic assays described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of aberrant 25943 expression or activity, such that a cellular proliferation disorder is prevented or, alternatively, delayed in its progression. Depending on the type of 25943 aberrancy, for example, a 25943 molecule, 25943 agonist or 25943 antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

B. Therapeutic Methods

Another aspect of the invention pertains to methods for treating a subject suffering from a cellular proliferation disorder. These methods involve administering to a subject an agent which modulates 25943 expression or activity (*e.g.*, an agent identified by a screening assay described herein), or a combination of such agents. In another embodiment, the method involves administering to a subject a 25943 protein or nucleic acid molecule as therapy to compensate for reduced, aberrant, or unwanted 25943 expression or activity.

Stimulation of 25943 activity is desirable in situations in which 25943 is abnormally downregulated and/or in which increased 25943 activity is likely to have a beneficial effect, *i.e.*, an increase in cellular proliferation, thereby ameliorating a cellular proliferation disorder such as a neurodegenerative disorder in a subject. Likewise, inhibition of 25943 activity is desirable in situations in which 25943 is abnormally upregulated and/or in which decreased 25943 activity is likely to have a beneficial effect, *e.g.*, an decrease in cellular proliferation, thereby ameliorating a cellular proliferation disorder such as breast cancer, ovarian cancer, lung cancer, or colon cancer in a subject.

The agents which modulate 25943 activity can be administered to a subject using pharmaceutical compositions suitable for such administration. Such compositions

typically comprise the agent (*e.g.*, nucleic acid molecule, protein, or antibody) and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition used in the therapeutic methods of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyetheylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of

dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, and sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the agent that modulates 25943 activity (*e.g.*, a fragment of a 25943 protein or an anti-25943 antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The agents that modulate 25943 activity can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the agents that modulate 25943 activity are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the agent that modulates 25943 activity and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an agent for the treatment of subjects.

Toxicity and therapeutic efficacy of such agents can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is

the therapeutic index and can be expressed as the ratio LD50/ED50. Agents which exhibit large therapeutic indices are preferred. While agents that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such agents to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such 25943 modulating agents lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any agent used in the therapeutic methods of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

As defined herein, a therapeutically effective amount of protein or polypeptide (*i.e.*, an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, preferably about 0.01 to 25 mg/kg body weight, more preferably about 0.1 to 20 mg/kg body weight, and even more preferably about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight. The skilled artisan will appreciate that certain factors may influence the dosage required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments.

In a preferred example, a subject is treated with antibody, protein, or polypeptide in the range of between about 0.1 to 20 mg/kg body weight, one time per week for between about 1 to 10 weeks, preferably between 2 to 8 weeks, more preferably between about 3 to 7 weeks, and even more preferably for about 4, 5, or 6 weeks. It will also be appreciated that the effective dosage of antibody, protein, or polypeptide used for treatment may increase or decrease over the course of a particular treatment. Changes in dosage may result and become apparent from the results of diagnostic assays as described herein.

The present invention encompasses agents which modulate expression or activity. An agent may, for example, be a small molecule. For example, such small molecules include, but are not limited to, peptides, peptidomimetics, amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (*i.e.*, including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds. It is understood that appropriate doses of small molecule agents depends upon a number of factors within the ken of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of the small molecule will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the small molecule to have upon the nucleic acid or polypeptide of the invention.

Exemplary doses include milligram or microgram amounts of the small molecule per kilogram of subject or sample weight (*e.g.*, about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram). It is furthermore understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. Such appropriate doses may be determined using the assays described herein. When one or more of these small molecules is to be administered to an animal (*e.g.*, a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

Further, an antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (*e.g.*, methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (*e.g.*, mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclothosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (*e.g.*, daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (*e.g.*, dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (*e.g.*, vincristine and vinblastine).

The conjugates of the invention can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, *e.g.*, Arnon *et al.*, "Monoclonal Antibodies for Immunotargeting of Drugs in Cancer Therapy", in *Monoclonal Antibodies and Cancer Therapy*, Reisfeld *et al.* (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom *et al.*, "Antibodies for Drug Delivery", in *Controlled Drug Delivery* (2nd Ed.), Robinson *et al.* (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers of Cytotoxic Agents in Cancer Therapy: A Review", in *Monoclonal Antibodies '84: Biological and Clinical Applications*, Pinchera *et al.* (eds.), pp. 475-506 (1985); "Analysis, Results, and Future Prospective of the Therapeutic Use of Radiolabeled Antibody in Cancer Therapy", in *Monoclonal Antibodies*

for Cancer Detection and Therapy, Baldwin *et al.* (eds.), pp. 303-16 (Academic Press 1985), and Thorpe *et al.* (1982) "The Preparation and Cytotoxic Properties of Antibody-Toxin Conjugates", *Immunol. Rev.* 62:119-58. Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

The nucleic acid molecules used in the methods of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see U.S. Patent No. 5,328,470) or by stereotactic injection (see, *e.g.*, Chen *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

C. Pharmacogenomics

In conjunction with the therapeutic methods of the invention, pharmacogenomics (*i.e.*, the study of the relationship between a subject's genotype and that subject's response to a foreign compound or drug) may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, a physician or clinician may consider applying knowledge obtained in relevant pharmacogenomics studies in determining whether to administer an agent which modulates 25943 activity, as well as tailoring the dosage and/or therapeutic regimen of treatment with an agent which modulates 25943 activity.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, for example, Eichelbaum, M. *et al.* (1996) *Clin. Exp. Pharmacol. Physiol.* 23(10-11):983-985 and Linder, M.W. *et al.* (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare genetic

defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate aminopeptidase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

One pharmacogenomics approach to identifying genes that predict drug response, known as "a genome-wide association", relies primarily on a high-resolution map of the human genome consisting of already known gene-related markers (*e.g.*, a "bi-allelic" gene marker map which consists of 60,000-100,000 polymorphic or variable sites on the human genome, each of which has two variants). Such a high-resolution genetic map can be compared to a map of the genome of each of a statistically significant number of patients taking part in a Phase II/III drug trial to identify markers associated with a particular observed drug response or side effect. Alternatively, such a high resolution map can be generated from a combination of some ten million known single nucleotide polymorphisms (SNPs) in the human genome. As used herein, a "SNP" is a common alteration that occurs in a single nucleotide base in a stretch of DNA. For example, a SNP may occur once per every 1000 bases of DNA. A SNP may be involved in a disease process, however, the vast majority may not be disease-associated. Given a genetic map based on the occurrence of such SNPs, individuals can be grouped into genetic categories depending on a particular pattern of SNPs in their individual genome. In such a manner, treatment regimens can be tailored to groups of genetically similar individuals, taking into account traits that may be common among such genetically similar individuals.

Alternatively, a method termed the "candidate gene approach" can be utilized to identify genes that predict drug response. According to this method, if a gene that encodes a drug target is known (*e.g.*, a 25943 protein of the present invention), all common variants of that gene can be fairly easily identified in the population and it can be determined if having one version of the gene versus another is associated with a particular drug response.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and the cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among

different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Alternatively, a method termed the "gene expression profiling" can be utilized to identify genes that predict drug response. For example, the gene expression of an animal dosed with a drug (*e.g.*, a 25943 molecule or 25943 modulator of the present invention) can give an indication whether gene pathways related to toxicity have been turned on.

Information generated from more than one of the above pharmacogenomics approaches can be used to determine appropriate dosage and treatment regimens for prophylactic or therapeutic treatment of a subject. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and, thus, enhance therapeutic or prophylactic efficiency when treating a subject suffering from a cellular proliferation disorder with an agent which modulates 25943 activity.

IV. Recombinant Expression Vectors and Host Cells Used in the Methods of the Invention.

The methods of the invention (*e.g.*, the screening assays described herein) include the use of vectors, preferably expression vectors, containing a nucleic acid encoding a 25943 protein (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of

genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors to be used in the methods of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel (1990) *Methods Enzymol.* 185:3-7. Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cells and those which direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, 25943 proteins, mutant forms of 25943 proteins, fusion proteins, and the like).

The recombinant expression vectors to be used in the methods of the invention can be designed for expression of 25943 proteins in prokaryotic or eukaryotic cells. For example, 25943 proteins can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors), yeast cells, or mammalian cells. Suitable host cells are discussed further in Goeddel (1990) *supra*. Alternatively, the recombinant expression

vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Purified fusion proteins can be utilized in 25943 activity assays, (*e.g.*, direct assays or competitive assays described in detail below), or to generate antibodies specific for 25943 proteins. In a preferred embodiment, a 25943 fusion protein expressed in a retroviral expression vector of the present invention can be utilized to infect bone marrow cells which are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (*e.g.*, six weeks).

In another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J. *et al.*, *Molecular Cloning: A Laboratory Manual. 2nd ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

The methods of the invention may further use a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to 25943 mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific, or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid, or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes, see Weintraub, H. *et al.*, Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to the use of host cells into which a 25943 nucleic acid molecule of the invention is introduced, *e.g.*, a 25943 nucleic acid molecule within a recombinant expression vector or a 25943 nucleic acid molecule containing

sequences which allow it to homologously recombine into a specific site of the host cell's genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, a 25943 protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook *et al.* (*Molecular Cloning: A Laboratory Manual. 2nd ed., Cold Spring Harbor Laboratory*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

A host cell used in the methods of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) a 25943 protein. Accordingly, the invention further provides methods for producing a 25943 protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of the invention (into which a recombinant expression vector encoding a 25943 protein has been introduced) in a suitable medium such that a 25943 protein is produced. In another embodiment, the method further comprises isolating a 25943 protein from the medium or the host cell.

V. Isolated Nucleic Acid Molecules Used in the Methods of the Invention

The cDNA sequence of the isolated human 25943 gene and the predicted amino acid sequence of the human 25943 polypeptide are shown in SEQ ID NOs:1 and 2, respectively. Nucleotides 123-1046 of SEQ ID NO:1, set forth as SEQ ID NO:3, comprise the 25943 coding region.

The methods of the invention include the use of isolated nucleic acid molecules that encode 25943 proteins or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes to identify 25943-encoding nucleic acid molecules (*e.g.*, 25943 mRNA) and fragments for use as PCR primers for the amplification or mutation of 25943 nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

A nucleic acid molecule used in the methods of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or portion of the nucleic acid sequence of SEQ ID NO:1 as a hybridization probe, 25943 nucleic acid molecules can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, J. *et al.*, *Molecular Cloning: A Laboratory Manual. 2nd ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Moreover, a nucleic acid molecule encompassing all or a portion of SEQ ID NO:1 can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence of SEQ ID NO:1.

A nucleic acid used in the methods of the invention can be amplified using cDNA, mRNA or, alternatively, genomic DNA as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. Furthermore, oligonucleotides corresponding to 25943 nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

In a preferred embodiment, the isolated nucleic acid molecules used in the methods of the invention comprise the nucleotide sequence shown in SEQ ID NO:1, a complement of the nucleotide sequence shown in SEQ ID NO:1, or a portion of any of these nucleotide sequences. A nucleic acid molecule which is complementary to the nucleotide sequence shown in SEQ ID NO:1, is one which is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1 such that it can hybridize to the nucleotide sequence shown in SEQ ID NO:1 thereby forming a stable duplex.

In still another preferred embodiment, an isolated nucleic acid molecule used in the methods of the present invention comprises a nucleotide sequence which is at least about

55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9% or more identical to the entire length of the nucleotide sequence shown in SEQ ID NO:1, or a portion of any of this nucleotide sequence.

Moreover, the nucleic acid molecules used in the methods of the invention can comprise only a portion of the nucleic acid sequence of SEQ ID NO:1, for example, a fragment which can be used as a probe or primer or a fragment encoding a portion of a 25943 protein, *e.g.*, a biologically active portion of a 25943 protein. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12 or 15, preferably about 20 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, or 75 consecutive nucleotides of a sense sequence of SEQ ID NO:1 or an anti-sense sequence of SEQ ID NO:1, or of a naturally occurring allelic variant or mutant of SEQ ID NO:1. In one embodiment, a nucleic acid molecule used in the methods of the present invention comprises a nucleotide sequence which is greater than 50, 100, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350 or more nucleotides in length and hybridizes under stringent hybridization conditions to a nucleic acid molecule of SEQ ID NO:1.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences that are significantly identical or homologous to each other remain hybridized to each other. Preferably, the conditions are such that sequences at least about 70%, more preferably at least about 80%, even more preferably at least about 85% or 90% identical to each other remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, Ausubel *et al.*, eds., John Wiley & Sons, Inc. (1995), sections 2, 4 and 6. Additional stringent conditions can be found in *Molecular Cloning: A Laboratory Manual*, Sambrook *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), chapters 7, 9 and 11. A preferred, non-limiting example of stringent hybridization conditions includes hybridization in 4X or 6X sodium chloride/sodium citrate (SSC), at about 65-70°C (or hybridization in 4X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 1X SSC, at about 65-70°C. A further preferred, non-limiting example of stringent hybridization conditions includes hybridization at 6X SSC at 45°C, followed by one or more washes in 0.2X SSC,

0.1% SDS at 65°C. A preferred, non-limiting example of highly stringent hybridization conditions includes hybridization in 1X SSC, at about 65-70°C (or hybridization in 1X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 0.3X SSC, at about 65-70°C. A preferred, non-limiting example of reduced stringency hybridization conditions includes hybridization in 4X or 6X SSC, at about 50-60°C (or alternatively hybridization in 6X SSC plus 50% formamide at about 40-45°C) followed by one or more washes in 2X SSC, at about 50-60°C. Ranges intermediate to the above-recited values, *e.g.*, at 65-70°C or at 42-50°C are also intended to be encompassed by the present invention. SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes each after hybridization is complete. The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, $T_m(°C) = 2(\# \text{ of } A + T \text{ bases}) + 4(\# \text{ of } G + C \text{ bases})$. For hybrids between 18 and 49 base pairs in length, $T_m(°C) = 81.5 + 16.6(\log_{10}[Na^+]) + 0.41(\%G+C) - (600/N)$, where N is the number of bases in the hybrid, and $[Na^+]$ is the concentration of sodium ions in the hybridization buffer ($[Na^+]$ for 1xSSC = 0.165 M). It will also be recognized by the skilled practitioner that additional reagents may be added to hybridization and/or wash buffers to decrease non-specific hybridization of nucleic acid molecules to membranes, for example, nitrocellulose or nylon membranes, including but not limited to blocking agents (*e.g.*, BSA or salmon or herring sperm carrier DNA), detergents (*e.g.*, SDS), chelating agents (*e.g.*, EDTA), Ficoll, PVP and the like. When using nylon membranes, in particular, an additional preferred, non-limiting example of stringent hybridization conditions is hybridization in 0.25-0.5M NaH_2PO_4, 7% SDS at about 65°C, followed by one or more washes at 0.02M NaH_2PO_4, 1% SDS at 65°C, see *e.g.*, Church and Gilbert (1984) *Proc. Natl. Acad. Sci. USA* 81:1991-1995, (or alternatively 0.2X SSC, 1% SDS).

In preferred embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a 25943 protein, such as by measuring a level

of a 25943-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting 25943 mRNA levels or determining whether a genomic 25943 gene has been mutated or deleted.

The methods of the invention further encompass the use of nucleic acid molecules that differ from the nucleotide sequence shown in SEQ ID NO:1 due to degeneracy of the genetic code and thus encode the same 25943 proteins as those encoded by the nucleotide sequence shown in SEQ ID NO:1. In another embodiment, an isolated nucleic acid molecule included in the methods of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ ID NO:2.

The methods of the invention further include the use of allelic variants of human 25943, *e.g.*, functional and non-functional allelic variants. Functional allelic variants are naturally occurring amino acid sequence variants of the human 25943 protein that maintain a 25943 activity. Functional allelic variants will typically contain only conservative substitution of one or more amino acids of SEQ ID NO:2, or substitution, deletion or insertion of non-critical residues in non-critical regions of the protein. Non-functional allelic variants are naturally occurring amino acid sequence variants of the human 25943 protein that do not have a 25943 activity. Non-functional allelic variants will typically contain a non-conservative substitution, deletion, or insertion or premature truncation of the amino acid sequence of SEQ ID NO:2, or a substitution, insertion or deletion in critical residues or critical regions of the protein.

The methods of the present invention may further use non-human orthologues of the human 25943 protein. Orthologues of the human 25943 protein are proteins that are isolated from non-human organisms and possess the same 25943 activity.

The methods of the present invention further include the use of nucleic acid molecules comprising the nucleotide sequence of SEQ ID NO:1, or a portion thereof, in which a mutation has been introduced. The mutation may lead to amino acid substitutions at "non-essential" amino acid residues or at "essential" amino acid residues. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of 25943 (*e.g.*, the sequence of SEQ ID NO:2) without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the 25943 proteins of the present invention and other members of the glycosylasparaginase family are not likely to be amenable to alteration.

Mutations can be introduced into SEQ ID NO:1 by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino

acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, glycine, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a 25943 protein is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a 25943 coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for 25943 biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1, the encoded protein can be expressed recombinantly and the activity of the protein can be determined using an assay described herein.

Another aspect of the invention pertains to the use of isolated nucleic acid molecules which are antisense to the nucleotide sequence of SEQ ID NO:1. An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire 25943 coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding a 25943. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding 25943. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding 25943 disclosed herein, antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick

base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of 25943 mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of 25943 mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of 25943 mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules used in the methods of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a 25943 protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex,

or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention include direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule used in the methods of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett.* 215:327-330).

In still another embodiment, an antisense nucleic acid used in the methods of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haseloff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave 25943 mRNA transcripts to thereby inhibit translation of 25943 mRNA. A ribozyme having specificity for a 25943-encoding nucleic acid can be designed based upon the nucleotide sequence of a 25943 cDNA disclosed herein (*i.e.*, SEQ ID NO:1). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a 25943-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, 25943 mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

Alternatively, 25943 gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the 25943 (*e.g.*, the 25943 promoter and/or enhancers) to form triple helical structures that prevent transcription of the 25943 gene in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. (1992) *Bioessays* 14(12):807-15.

In yet another embodiment, the 25943 nucleic acid molecules used in the methods of the present invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup, B. and Nielsen, P.E. (1996) *Bioorg. Med. Chem.* 4(1):5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup B. and Nielsen (1996) *supra* and Perry-O'Keefe *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93:14670-675.

PNAs of 25943 nucleic acid molecules can be used in the therapeutic and diagnostic applications described herein. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNAs of 25943 nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (*e.g.*, by PNA-directed PCR clamping); as 'artificial restriction enzymes' when used in combination with other enzymes, (*e.g.*, S1 nucleases (Hyrup and Nielsen (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup and Nielsen (1996) *supra*; Perry-O'Keefe *et al.* (1996) *supra*).

In another embodiment, PNAs of 25943 can be modified, (*e.g.*, to enhance their stability or cellular uptake), by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of 25943 nucleic acid molecules can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, (*e.g.*, RNAse H and DNA

polymerases), to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup and Nielsen (1996) *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup and Nielsen (1996) *supra* and Finn P.J. *et al.* (1996) *Nucleic Acids Res.* 24 (17):3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used as a between the PNA and the 5' end of DNA (Mag, M. *et al.* (1989) *Nucleic Acids Res.* 17:5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) *supra*). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser, K.H. *et al.* (1975) *Bioorganic Med. Chem. Lett.* 5:1119-11124).

In other embodiments, the oligonucleotide used in the methods of the invention may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. WO 88/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. WO 89/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (See, *e.g.*, Krol *et al.* (1988) *Biotechniques* 6:958-976) or intercalating agents (see, *e.g.*, Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, (*e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

VI. Isolated 25943 Proteins and Anti-25943 Antibodies Used in the Methods of the Invention

The methods of the invention include the use of isolated 25943 proteins, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise anti-25943 antibodies. In one embodiment, native 25943 proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, 25943 proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a

25943 protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

As used herein, a "biologically active portion" of a 25943 protein includes a fragment of a 25943 protein having a 25943 activity. Biologically active portions of a 25943 protein include peptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the 25943 protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:2, which include fewer amino acids than the full length 25943 proteins, and exhibit at least one activity of a 25943 protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the 25943 protein. A biologically active portion of a 25943 protein can be a polypeptide which is, for example, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300 or more amino acids in length. Biologically active portions of a 25943 protein can be used as targets for developing agents which modulate a 25943 activity.

In a preferred embodiment, the 25943 protein used in the methods of the invention has an amino acid sequence shown in SEQ ID NO:2. In other embodiments, the 25943 protein is substantially identical to SEQ ID NO:2, and retains the functional activity of the protein of SEQ ID NO:2, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail in subsection V above. Accordingly, in another embodiment, the 25943 protein used in the methods of the invention is a protein which comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 76%, 77%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9% or more identical to SEQ ID NO:2.

To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-identical sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, or 90% of the length of the reference sequence (*e.g.*, when aligning a second sequence to the 25943 amino acid sequence of SEQ ID NO:2 having 308 amino acid residues, at least 92, preferably at least 123, more preferably at least 154, even more preferably at least 185, and even more preferably at least 216, 246, 277 or more amino acid residues are aligned). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide

positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* 48:444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at http://www.gcg.com), using either a Blosum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at http://www.gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of Meyers, E. and Miller, W. (*Comput. Appl. Biosci.* 4:11-17 (1988)) which has been incorporated into the ALIGN program (version 2.0 or 2.0U), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The methods of the invention may also use 25943 chimeric or fusion proteins. As used herein, a 25943 "chimeric protein" or "fusion protein" comprises a 25943 polypeptide operatively linked to a non-25943 polypeptide. A "25943 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a 25943 molecule, whereas a "non-25943 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the 25943 protein, *e.g.*, a protein which is different from the 25943 protein and which is derived from the same or a different organism. Within a 25943 fusion protein the 25943 polypeptide can correspond to all or a portion of a 25943 protein. In a preferred embodiment, a 25943 fusion protein comprises at least one biologically active portion of a 25943 protein. In another preferred embodiment, a 25943 fusion protein comprises at least two biologically active portions of a 25943 protein. Within the fusion protein, the term "operatively

linked" is intended to indicate that the 25943 polypeptide and the non-25943 polypeptide are fused in-frame to each other. The non-25943 polypeptide can be fused to the N-terminus or C-terminus of the 25943 polypeptide.

For example, in one embodiment, the fusion protein is a GST-25943 fusion protein in which the 25943 sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant 25943.

In another embodiment, this fusion protein is a 25943 protein containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of 25943 can be increased through use of a heterologous signal sequence.

The 25943 fusion proteins used in the methods of the invention can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. The 25943 fusion proteins can be used to affect the bioavailability of a 25943 substrate. Use of 25943 fusion proteins may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a 25943 protein; (ii) mis-regulation of the 25943 gene; and (iii) aberrant post-translational modification of a 25943 protein.

Moreover, the 25943-fusion proteins used in the methods of the invention can be used as immunogens to produce anti-25943 antibodies in a subject, to purify 25943 ligands and in screening assays to identify molecules which inhibit the interaction of 25943 with a 25943 substrate.

Preferably, a 25943 chimeric or fusion protein used in the methods of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons:1992). Moreover, many expression vectors are commercially

available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A 25943-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the 25943 protein.

The present invention also pertains to the use of variants of the 25943 proteins which function as either 25943 agonists (mimetics) or as 25943 antagonists. Variants of the 25943 proteins can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of a 25943 protein. An agonist of the 25943 proteins can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of a 25943 protein. An antagonist of a 25943 protein can inhibit one or more of the activities of the naturally occurring form of the 25943 protein by, for example, competitively modulating a 25943-mediated activity of a 25943 protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the 25943 protein.

In one embodiment, variants of a 25943 protein which function as either 25943 agonists (mimetics) or as 25943 antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of a 25943 protein for 25943 protein agonist or antagonist activity. In one embodiment, a variegated library of 25943 variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of 25943 variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential 25943 sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of 25943 sequences therein. There are a variety of methods which can be used to produce libraries of potential 25943 variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential 25943 sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucleic Acids Res.* 11:477).

In addition, libraries of fragments of a 25943 protein coding sequence can be used to generate a variegated population of 25943 fragments for screening and subsequent selection of variants of a 25943 protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a 25943 coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the 25943 protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of 25943 proteins. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify 25943 variants (Arkin and Youvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delagrave *et al.* (1993) *Protein Eng.* 6(3):327-331).

The methods of the present invention further include the use of anti-25943 antibodies. An isolated 25943 protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind 25943 using standard techniques for polyclonal and monoclonal antibody preparation. A full-length 25943 protein can be used or, alternatively, antigenic peptide fragments of 25943 can be used as immunogens. The antigenic peptide of 25943 comprises at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:2 and encompasses an epitope of 25943 such that an antibody raised against the peptide forms a specific immune complex with the 25943 protein. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more

preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues.

Preferred epitopes encompassed by the antigenic peptide are regions of 25943 that are located on the surface of the protein, *e.g.*, hydrophilic regions, as well as regions with high antigenicity.

A 25943 immunogen is typically used to prepare antibodies by immunizing a suitable subject, (*e.g.*, rabbit, goat, mouse, or other mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly expressed 25943 protein or a chemically synthesized 25943 polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with an immunogenic 25943 preparation induces a polyclonal anti-25943 antibody response.

The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site which specifically binds (immunoreacts with) an antigen, such as a 25943. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and $F(ab')_2$ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies that bind 25943 molecules. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of 25943. A monoclonal antibody composition thus typically displays a single binding affinity for a particular 25943 protein with which it immunoreacts.

Polyclonal anti-25943 antibodies can be prepared as described above by immunizing a suitable subject with a 25943 immunogen. The anti-25943 antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized 25943. If desired, the antibody molecules directed against 25943 can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction. At an appropriate time after immunization, *e.g.*, when the anti-25943 antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497) (see also, Brown *et al.* (1981) *J. Immunol.* 127:539-46;

Brown *et al.* (1980) *J. Biol. Chem.* 255:4980-83; Yeh *et al.* (1976) *Proc. Natl. Acad. Sci. USA* 76:2927-31; and Yeh *et al.* (1982) *Int. J. Cancer* 29:269-75), the more recent human B cell hybridoma technique (Kozbor *et al.* (1983) *Immunol. Today* 4:72), the EBV-hybridoma technique (Cole *et al.* (1985) *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing monoclonal antibody hybridomas is well known (see generally Kenneth, R.H. in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, New York (1980); Lerner, E.A. (1981) *Yale J. Biol. Med.* 54:387-402; Gefter, M.L. *et al.* (1977) *Somat. Cell Genet.* 3:231-36). Briefly, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with a 25943 immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds 25943.

Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti-25943 monoclonal antibody (see, *e.g.*, Galfre, G. *et al.* (1977) *Nature* 266:55052; Gefter *et al.* (1977) *supra*; Lerner (1981) *supra*; and Kenneth (1980) *supra*). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods which also would be useful. Typically, the immortal cell line (*e.g.*, a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing hypoxanthine, aminopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, *e.g.*, the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/O-Ag14 myeloma lines. These myeloma lines are available from ATCC. Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind 25943, *e.g.*, using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-25943 antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phage display library) with 25943 to thereby isolate immunoglobulin library members that bind 25943. Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP™ Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, Ladner *et al.* U.S. Patent No. 5,223,409; Kang *et al.* PCT International Publication No. WO 92/18619; Dower *et al.* PCT International Publication No. WO 91/17271; Winter *et al.* PCT International Publication WO 92/20791; Markland *et al.* PCT International Publication No. WO 92/15679; Breitling *et al.* PCT International Publication WO 93/01288; McCafferty *et al.* PCT International Publication No. WO 92/01047; Garrard *et al.* PCT International Publication No. WO 92/09690; Ladner *et al.* PCT International Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1369-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J.* 12:725-734; Hawkins *et al.* (1992) *J. Mol. Biol.* 226:889-896; Clackson *et al.* (1991) *Nature* 352:624-628; Gram *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3576-3580; Garrard *et al.* (1991) *Biotechnology (NY)* 9:1373-1377; Hoogenboom *et al.* (1991) *Nucleic Acids Res.* 19:4133-4137; Barbas *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:7978-7982; and McCafferty *et al.* (1990) *Nature* 348:552-554.

Additionally, recombinant anti-25943 antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the methods of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson *et al.* International Application No. PCT/US86/02269; Akira, *et al.* European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison *et al.* European Patent Application 173,494; Neuberger *et al.* PCT International Publication No. WO 86/01533; Cabilly *et al.* U.S. Patent No. 4,816,567; Cabilly *et al.* European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521-3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.*

(1987) *Cancer Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559; Morrison, S.L. (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; Winter U.S. Patent No. 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyen *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

An anti-25943 antibody can be used to detect 25943 protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the 25943 protein. Anti-25943 antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

VII. Electronic Apparatus Readable Media and Arrays

Electronic apparatus readable media comprising 25943 sequence information is also provided. As used herein, "25943 sequence information" refers to any nucleotide and/or amino acid sequence information particular to the 25943 molecules of the present invention, including but not limited to full-length nucleotide and/or amino acid sequences, partial nucleotide and/or amino acid sequences, polymorphic sequences including single nucleotide polymorphisms (SNPs), epitope sequences, and the like. Moreover, information "related to" said 25943 sequence information includes detection of the presence or absence of a sequence (*e.g.*, detection of expression of a sequence, fragment, polymorphism, etc.), determination of the level of a sequence (*e.g.*, detection of a level of expression, for example, a quantitative detection), detection of a reactivity to a sequence (*e.g.*, detection of protein expression and/or levels, for example, using a sequence-specific

antibody), and the like. As used herein, "electronic apparatus readable media" refers to any suitable medium for storing, holding, or containing data or information that can be read and accessed directly by an electronic apparatus. Such media can include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as compact discs; electronic storage media such as RAM, ROM, EPROM, EEPROM and the like; and general hard disks and hybrids of these categories such as magnetic/optical storage media. The medium is adapted or configured for having recorded thereon 25943 sequence information of the present invention.

As used herein, the term "electronic apparatus" is intended to include any suitable computing or processing apparatus or other device configured or adapted for storing data or information. Examples of electronic apparatus suitable for use with the present invention include stand-alone computing apparatuses; networks, including a local area network (LAN), a wide area network (WAN) Internet, Intranet, and Extranet; electronic appliances such as a personal digital assistants (PDAs), cellular phone, pager and the like; and local and distributed processing systems.

As used herein, "recorded" refers to a process for storing or encoding information on the electronic apparatus readable medium. Those skilled in the art can readily adopt any of the presently known methods for recording information on known media to generate manufactures comprising the 25943 sequence information.

A variety of software programs and formats can be used to store the sequence information on the electronic apparatus readable medium. For example, the sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, represented in the form of an ASCII file, or stored in a database application, such as DB2, Sybase, Oracle, or the like, as well as in other forms. Any number of dataprocessor structuring formats (*e.g.*, text file or database) may be employed in order to obtain or create a medium having recorded thereon the 25943 sequence information.

By providing 25943 sequence information in readable form, one can routinely access the sequence information for a variety of purposes. For example, one skilled in the art can use the sequence information in readable form to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequences of the invention which match a particular target sequence or target motif.

The present invention therefore provides a medium for holding instructions for performing a method for determining whether a subject has a 25943 associated disease or disorder or a pre-disposition to a 25943 associated disease or disorder, wherein the method comprises the steps of determining 25943 sequence information associated with the subject and based on the 25943 sequence information, determining whether the subject has a 25943 associated disease or disorder or a pre-disposition to a 25943 associated disease or disorder, and/or recommending a particular treatment for the disease, disorder, or pre-disease condition.

The present invention further provides in an electronic system and/or in a network, a method for determining whether a subject has a 25943 associated disease or disorder or a pre-disposition to a disease associated with 25943 wherein the method comprises the steps of determining 25943 sequence information associated with the subject, and based on the 25943 sequence information, determining whether the subject has a 25943 associated disease or disorder or a pre-disposition to a 25943 associated disease or disorder, and/or recommending a particular treatment for the disease, disorder or pre-disease condition. The method may further comprise the step of receiving phenotypic information associated with the subject and/or acquiring from a network phenotypic information associated with the subject.

The present invention also provides in a network, a method for determining whether a subject has a 25943 associated disease or disorder or a pre-disposition to a 25943 associated disease or disorder associated with 25943, said method comprising the steps of receiving 25943 sequence information from the subject and/or information related thereto, receiving phenotypic information associated with the subject, acquiring information from the network corresponding to 25943 and/or a 25943 associated disease or disorder, and based on one or more of the phenotypic information, the 25943 information (*e.g.*, sequence information and/or information related thereto), and the acquired information, determining whether the subject has a 25943 associated disease or disorder or a pre-disposition to a 25943 associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder or pre-disease condition.

The present invention also provides a business method for determining whether a subject has a 25943 associated disease or disorder or a pre-disposition to a 25943 associated disease or disorder, said method comprising the steps of receiving information related to 25943 (*e.g.*, sequence information and/or information related thereto), receiving

phenotypic information associated with the subject, acquiring information from the network related to 25943 and/or related to a 25943 associated disease or disorder, and based on one or more of the phenotypic information, the 25943 information, and the acquired information, determining whether the subject has a 25943 associated disease or disorder or a pre-disposition to a 25943 associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder or pre-disease condition.

The invention also includes an array comprising a 25943 sequence of the present invention. The array can be used to assay expression of one or more genes in the array. In one embodiment, the array can be used to assay gene expression in a tissue to ascertain tissue specificity of genes in the array. In this manner, up to about 7600 genes can be simultaneously assayed for expression, one of which can be 25943. This allows a profile to be developed showing a battery of genes specifically expressed in one or more tissues.

In addition to such qualitative determination, the invention allows the quantitation of gene expression. Thus, not only tissue specificity, but also the level of expression of a battery of genes in the tissue is ascertainable. Thus, genes can be grouped on the basis of their tissue expression *per se* and level of expression in that tissue. This is useful, for example, in ascertaining the relationship of gene expression between or among tissues. Thus, one tissue can be perturbed and the effect on gene expression in a second tissue can be determined. In this context, the effect of one cell type on another cell type in response to a biological stimulus can be determined. Such a determination is useful, for example, to know the effect of cell-cell interaction at the level of gene expression. If an agent is administered therapeutically to treat one cell type but has an undesirable effect on another cell type, the invention provides an assay to determine the molecular basis of the undesirable effect and thus provides the opportunity to co-administer a counteracting agent or otherwise treat the undesired effect. Similarly, even within a single cell type, undesirable biological effects can be determined at the molecular level. Thus, the effects of an agent on expression of other than the target gene can be ascertained and counteracted.

In another embodiment, the array can be used to monitor the time course of expression of one or more genes in the array. This can occur in various biological contexts, as disclosed herein, for example development of a 25943 associated disease or disorder, progression of 25943 associated disease or disorder, and processes, such a cellular transformation associated with the 25943 associated disease or disorder.

The array is also useful for ascertaining the effect of the expression of a gene on the expression of other genes in the same cell or in different cells (*e.g.*, ascertaining the effect of 25943 expression on the expression of other genes). This provides, for example, for a selection of alternate molecular targets for therapeutic intervention if the ultimate or downstream target cannot be regulated.

The array is also useful for ascertaining differential expression patterns of one or more genes in normal and abnormal cells. This provides a battery of genes (*e.g.*, including 25943) that could serve as a molecular target for diagnosis or therapeutic intervention.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application, are incorporated herein by reference.

EXAMPLES

EXAMPLE 1: MEASUREMENT OF GLYCOSYLASPARAGINASE ACTIVITY (METHOD 1)

This example describes methods for determining the activity of a glycosylasparaginase molecule, *e.g.*, a 25943 molecule. The methods are performed according to Noronkoski, T. *et al.* (1998) *J. Biol. Chem.* 273:26295-26297, incorporated herein by reference.

Materials

25943 may purified to homogeneity from 25943 expressing cells according to the methods of Mononen, I. *et al.* (1995) *FASEB J.* 9:428-433 and Kaartinen, V. *et al.* (1991) *J. Biol. Chem.* 266:5860-5869, incorporated herein by reference. The purification protocol includes caprylic acid precipitation, affinity chromatography with concanavalin A lectin, gel filtration, hydrophobic interaction chromatography, and anion exchange chromatography. The fractions containing 25943 activity are pooled, and the purity of the preparation is estimated using SDS-polyacrylamide gel electrophoresis and silver staining, using standard methods. β-Aspartame and β-aspartylglycine may be purchased from Bachem Feinchemikalien AG (Bubendorf, Switzerland). Other peptides are prepared as described in Hanson, R.W. and Rydon, H.N. (1964) *J. Chem. Soc.* 115:836-842 and Cohen-Anisfeld, S.T. and Lansbury, P.T., Jr. (1993) *J. Am. Chem. Soc.* 115:10531-10537,

incorporated herein by reference. H-Ser-HN_2 and other amides are created as described in Hanson and Rydon (1964) *supra* and Fastrez, J. and Fersht, A.R. (1973) *Biochemistry* 12:2025-2034, incorporated herein by reference. β-Glycosylamine (1-amino-N-acetylglucosamine; GlcNAc-NH_2), β-aspartylglucosamine (GlcNAc-Asn), and aspartic acid β-methyl ester (H-Asp(OMe)-OH) are from Sigma (St. Louis, MO).

Enzyme Assays

Glycosylasparaginase activity of 25943 is measured with a fluorometric method using AspAMC as a substrate (Mononen, I. *et al.* (1993) *Anal. Biochem.* 208:372-374, incorporated herein by reference). The kinetic parameters for glycosylasparaginase catalyzed hydrolysis of β-aspartyl compounds are determined with a spectrophotometric assay (Tarantino, A.L. and Maley, F. (1969) *Arch. Biochem. Biophys.* 130:295-303; Noronkoski, T. and Mononen, I. (1997) *Glycobiology* 7:217-220). The least square Lineweaver-Burk analysis is used in the determination of Michaelis constant (K_m) and maximum reaction velocity (V_{max}).

Assay of β-Aspartyl Peptides

Formation of β-aspartyl peptides during their 25943 catalyzed synthesis is measured as described in Mononen, I. *et al.* (1996) *Biochem. Biophys. Res. Commun.* 218-510-513, incorporated herein by reference. Various amounts of β-aspartyl donors and β-aspartyl acceptors are incubated in the presence of 25943 in 50 mM Tris-HCl buffer, pH 7.5, at 37°C. Aliquots of the incubation mixture are injected onto an HPLC column, and the formation of β-aspartyl peptides is measured. High performance liquid chromatography is performed as described in Mononen *et al.* (1996) *supra* using a 250 X 4.6 mm (inner diameter) amino column (Spherisorb-NH_2, 5 μm particles), isocratic elution with 2.5 mM KH_2PO_4/aceotnitrile (40/60 v/v, pH 4.6), flow rate 1 ml/min and detection wavelength 214 nm.

EXAMPLE 2: MEASUREMENT OF GLYCOSYLASPARAGINASE ACTIVITY (METHOD 2)

This example describes methods for determining the activity of a glycosylasparaginase molecule, *e.g.*, a 25943 molecule. The methods are performed

according to Noronkoski, T. *et al.* (1997) *FEBS Lett.* 412:149-152, incorporated herein by reference.

Materials

25943 may be purified as described above in Example 4. Total protein is determined using a Bio-Rad protein assay kit (Bio-Rad laboratories, Hercules, CA, U.S.A.). Aspartylglucosamine, phenylisothiocyanate, and carboxymethyl cysteine are products of Sigma Chemical Co., St. Louis, MO, U.S.A. Asparagine and aspartic acid may be purchased from E. Merck, Darmstadt, Germany. All other reagents are of analytical grade and are used without further purification. High-performance liquid chromatography (HPLC) is carried out with a Merck/Hitachi L-6200 liquid chromatograph (Hitachi Ltd., Tokyo, Japan). The column is Spherisorb S3 ODS2 (150 X 4.6 mm, internal diameter) (Phase Separations Ltd., Deeside, U.K.).

Enzyme Assay

The assay for glycosylasparaginase activity is based on HPLC analysis of the reaction components GlcNAc-Asn, Asn and Asp after their phenylisothiocyanate (PITC) derivitization (Mononen, I.T. *et al.* (1993) *Anal. Biochem.* 208:372-374; Ebert, R.F. (1986) *Anal. Biochem.* 154:431-435; incorporated herein by reference). Carboxymethyl cysteine (CmCys) is used as an internal standard. The kinetic and inhibition constants are determined at 22°C, and the incubation mixture contains various amounts of GlcNAc-Asn and/or Asn, 0.8 mM CmCys, and 25943 in 50 mM sodium-potassium phosphate buffer, pH 7.5 in a total volume of 50 μl.

EXAMPLE 3: MEASUREMENT OF GLYCOSYLASPARAGINASE ACTIVITY (METHOD 3)

This example describes methods for determining the activity of a glycosylasparaginase molecule, *e.g.*, a 25943 molecule. The methods are performed according to Liu, Y. *et al.* (1996) *J. Biol. Chem.* 6:527-536, incorporated herein by reference.

Glycosylasparaginase activity is assayed by measuring the release of N-acetylglucosamine from the substrate N^4-(β-N-acetylglucosaminyl)-L-asparagine (GlcNAc-Asn) (Bachem, Torrance, CA, U.S.A.) (Tollersrud, O.K. and Aronson, N.N., Jr.

(1989) *Biochem. J.* 260:101-108). Reactions with 2.5 mM substrate in 20 µl of 20 mM sodium phosphate buffer, pH 7.5, are incubated with 25943 for appropriate times at 37°C and stopped by boiling for 3 minutes after adding 50 µl of 250 mM sodium borate buffer, pH 8.8. Released N-acetylglucosamine is assayed by the Morgan-Elson reaction. One unit of 25943 liberates 1 µmol of N-acetylglucosamine per minute.

EXAMPLE 4: PURIFICATION OF 25943 FROM Sf9 CELLS

The following methods are performed according to Liu, Y. *et al.* (1996) *J. Biol. Chem.* 6:527-536, incorporated herein by reference. 25943 is cloned into an appropriate vector and expressed in Sf9 insect cells. The cells are spin cultured in 10 liters of HyQ CCM-3 medium (Hyclone Lab. Inc., Logan, UT, U.S.A.) to a final cell density of 1.75×10^6/ml with a viability of 99%. Cells are collected at 4°C by centrifugation at 2000 rpm for 10 minutes. The cell pellet is washed once with phosphate buffered saline (PBS) and then resuspended in 50 ml of 50 mM sodium phosphate buffer, pH 7.5, containing 0.15 M NaCl. Cells are homogenized by sonication followed by centrifugation at 15,000 rpm for 30 minutes at 4°C. The cell lysate supernatant is subjected to chromatography in the cold on concanavalin A Sepharose (Sigma) equilibrated with 50 mM sodium phosphate buffer, pH 7.5, containing 0.15 M NaCl. Non-binding proteins are washed from the column using 300 ml buffer, and bound glycosylated proteins are eluted by 0.2 M methyl α-mannoside (Sigma). Eluted 25943 is concentrated to about 20 ml in an Amicon ultrafiltration cell fitted with a YM-10 membrane (Amicon Corp., Danvers, MA, U.S.A.). The sodium phosphate buffer is changed to 0.02 M Tris-HCl, pH 8.5, by three repeated ultrafiltrations. The concentrated 25943 sample is applied to a DE52-cellulose (Whatman Laboratory Division, Maidstone, England) anion-exchange column that has been pre-equilibrated with 0.02 M Tris-HCl, pH 8.5. Th column is eluted at 0.5 ml/minute with a 400 ml linear NaCl gradient (0-0.4 M in the same Tris buffer). Four ml fractions are collected, and those that contain 25943 are combined, concentrated to 3 ml and subjected to gel filtration on Sephadex G-150-120 (Sigma) equilibrated with 0.02 M sodium phosphate buffer, pH 5.5. The column is run at 0.25 ml/min and 2.5 ml fractions are collected. Fractions containing 25943 are combined and concentrated to 3 ml with at YM-10 membrane. This sample is subjected to CM52-cellulose (Whatman) cation-exchange chromatography on a column equilibrated with 0.02 M sodium phosphate buffer, pH 5.5. 25943 is eluted by 100 ml of the same buffer containing a linear NaCl gradient from 0 to

0.4 M. The column is run at 0.25 ml/minute and 2 ml fractions are collected. 25943 fractions are combined and concentrated to 1 ml by Centricon 10 ultrafiltration (Amicon Corp.).

EXAMPLE 5: SOFT AGAR ASSAY FOR ANCHORAGE-INDEPENDENT GROWTH OF CELLS

Base Agar

1% Agar (DNA grade) is melted in a microwave and cooled to 40°C in a waterbath. 2X RPMI medium + 20% fetal calf serum (FCS) is also warmed to 40°C in a waterbath. Equal volumes of the two solutions are mixed to give 0.5% Agar + 1X RPMI + 10% FCS. 1.5ml is poured into each 35 mm Petri plate and allowed to set. The plates can be stored at 4°C for up to 1 week.

Top Agar

0.7% Agarose (DNA grade) is melted in a microwave and cooled to 40°C in a waterbath. 2X RPMI + 20% FCS is also warmed to the same temperature.

Cells

The cells (*e.g.*, breast, ovary, lung, or colon cells) to be assayed are trypsinized, suspended in medium, and counted. A positive control, such as a ras transformed cell line, should always be used. The concentration of the cell suspension is adjusted to 200,000 cells/ml.

Plating and staining

0.1ml of cell suspension is added to 10 ml capped centrifuge tubes. The 35 mm Petri plates containing the base agar are removed from 4°C about 30 minutes prior to plating to allow them to warm up to room temperature. 3ml 2X RPMI + 10% or 20% FCS and 3 ml 0.7% Agarose are added to each tube of cell suspension and mixed gently. 1.5 ml of this mixture is added to each replicate plate (each plate is done in triplicate), and the agarose is allowed to solidify. The plates are incubated at 37°C in humidified incubator for 10 - 14 days. After completion of the incubation period, the plates are stained with 0.5 ml of 0.005% Crystal Violet for at least 1 hour. The colonies are then counted using a dissecting microscope.

EXAMPLE 6: FLUORIMETRIC ASSAY FOR GLYCOSYLASPARAGINASE ACTIVITY IN SERUM, PLASMA, OR LYMPHOCYTES

This example describes methods for determining the activity of a glycosylasparaginase molecule, *e.g.*, a 25943 molecule, from the serum, plasma, or lymphocytes of a subject. The methods are performed according to Mononen, I. *et al.* (1994) *Clin. Chem.* 40:385-388, incorporated herein by reference.

Materials

The substrate AspAMC and the standard 7-amino-4-methylcoumarin (AMC) may be purchased from Bachem AG, Bubendorf, Switzerland. Ficoll-Paque is purchased from Pharmacia, Uppsalla, Sweden. All other reagents are of analytical grade and are used without further purification.

Samples

Serum and plasma samples, as well as cultured fibroblast cell lines, are obtained from subjects. Lymphocytes are isolated by density gradient centrifugation with Ficoll-Paque according to the manufacturer's instructions. Samples are stored frozen at -20°C until analysis.

Effect of hemoglobin and bilirubin concentration

The presence of hemoglobin and/or bilirubin in the incubation mixture has a considerable inhibitory effect on the glycosylasparaginase activity. Accordingly, the measurements are compared to standards of hemoglobin and/or bilirubin concentrations.

Blood samples are hemolyzed by freezing 10 ml for 24 hours, and then thawing. Serum is separated after centrifugation at 3000 g for 10 minutes and mixed with normal serum in different ratios. The hemoglobin concentration is determined according to the method of Ferencz, A. and Basco, M. (1983) *Clin. Chem.* 134:103-106.

Glycosylasparaginase activity

The glycosylasparaginase activity of 25943 is assayed in plasma or serum by incubating 100 µl of AspAMC (10 mmol/L in ethylene glycol) and 90 µl of Tris-HCl buffer (50 mmol/L, pH 7.5) containing 10 ml/L ethylene glycol for 60-240 minutes at 37°C. Lymphocyte glycosylasparaginase activity is measured with 0.5 mmol/L substrate

in 50 mmol/L Tris-HCl, pH 7.5, containing 10 ml/L ethylene glycol in a final volume of 200 µl (Mononen, I.T., *et al.* (1993) *Anal. Biochem.* 208:372-374). Fluorescence is measured with an IL Multistat III Plus fluorescence centrifugal analyzer (Instrumentation Laboratory, Lexington, MA) at 350 nm (excitation) and 450 nm (emission). Glycosylasparaginase activity in plasma and serum is expressed as mU/L. The activity in lymphocytes is expressed as mU/g protein. One unit of glycosylasparaginase liberates 1 µmol of AMC from the substrate per minute at 37°C. Protein is determined with a protein assay kit (Bio-Rad Labs, Richmond, CA) and a Kone Compact clinical analyzer (Kone Instruments, Espoo, Finland).

EXAMPLE 7: FLUORIMETRIC ASSAY FOR GLYCOSYLASPARAGINASE ACTIVITY (METHOD 4)

This example describes methods for determining the activity of a glycosylasparaginase molecule, *e.g.*, a 25943 molecule or a 25943 modulator. The methods are performed according to Mononen, I.T. *et al.* (1994) *Analytical Biochem.* 208:372-374, incorporated herein by reference.

Materials

AspAMC and AMC are from Bachem AG (Bubendorf, Switzerland). 5-Diazo-4-oxo-L-norvaline, specific inhibitor of glycosylasparaginase (Kaartinen, V. *et al.* (1991) *J. Biol. Chem.* 266:5860-5869), is synthesized according to the methods of Handschumacher, R.E. *et al.* (1968) *Science* 161:62-63.

Biological Samples

Leukocytes are isolated from 10 ml of blood as described in Kaartinen, V. and Mononen, I. ((1990) *Anal. Biochem.* 190:98-101) and suspended, like fibroblasts isolated from tissue culture plates, in 50 mM Tris-HCl, pH 7.5, containing 0.2% Triton X-100 and 0.5 mM EDTA.

Assay

The fluorometric glycosylasparaginase assay is conducted at 37°C using 1 mM substrate in 50 mM Tris-HCl, pH 7.5, containing 0.2% ethylene glycol in a total volume of 100 µl. The release of 7-amino-4-methylcoumarin is measured fluorometrically using an

IL Multistat III plus fluorescence centrifugal analyzer (Instrumentation Laboratory, Lexington, MA). Excitation and emission wavelengths are 350 and 450 nm, respectively. For the calibration graph, the substrate is replaced by various concentrations of AMC. All the enzyme activity is expressed as μU/mg protein, where one unit indicates the enzyme amount causing the liberation of one micromole of AMC per minute at 37°C. Total protein concentration is determined using a Bio-Rad protein assay reagent kit.

EXAMPLE 8: TISSUE EXPRESSION ANALYSIS OF HUMAN 25943 mRNA USING TAQMAN ANALYSIS

This example describes the tissue distribution of human 25943 mRNA, as determined using the TaqMan™ procedure. The Taqman™ procedure is a quantitative, reverse transcription PCR-based approach for detecting mRNA. The RT-PCR reaction exploits the 5' nuclease activity of AmpliTaq Gold™ DNA Polymerase to cleave a TaqMan™ probe during PCR. Briefly, cDNA was generated from the samples of interest and used as the starting material for PCR amplification. In addition to the 5' and 3' gene-specific primers, a gene-specific oligonucleotide probe (complementary to the region being amplified) was included in the reaction (*i.e.*, the Taqman™ probe). The TaqMan™ probe included the oligonucleotide with a fluorescent reporter dye covalently linked to the 5' end of the probe (such as FAM (6-carboxyfluorescein), TET (6-carboxy-4,7,2',7'-tetrachlorofluorescein), JOE (6-carboxy-4,5-dichloro-2,7-dimethoxyfluorescein), or VIC) and a quencher dye (TAMRA (6-carboxy-N,N,N',N'-tetramethylrhodamine) at the 3' end of the probe.

During the PCR reaction, cleavage of the probe separated the reporter dye and the quencher dye, resulting in increased fluorescence of the reporter. Accumulation of PCR products was detected directly by monitoring the increase in fluorescence of the reporter dye. When the probe was intact, the proximity of the reporter dye to the quencher dye resulted in suppression of the reporter fluorescence. During PCR, if the target of interest is present, the probe specifically annealed between the forward and reverse primer sites. The 5'-3' nucleolytic activity of the AmpliTaq™ Gold DNA Polymerase cleaved the probe between the reporter and the quencher only if the probe hybridized to the target. The probe fragments were then displaced from the target, and polymerization of the strand continued. The 3' end of the probe was blocked to prevent extension of the probe during PCR. This process occurred in every cycle and did not interfere with the exponential

accumulation of product. RNA was prepared using the trizol method and treated with DNase to remove contaminating genomic DNA. cDNA was synthesized using standard techniques. Mock cDNA synthesis in the absence of reverse transcriptase resulted in samples with no detectable PCR amplification of the control gene confirms efficient removal of genomic DNA contamination.

The expression of human 25943 mRNA was examined in various cell types and tissues. As shown in Table 1, 25943 is highly expressed in human umbilical vein endothelial cells (HUVECs), kidney, normal skin, normal brain cortex, hypothalamus, ovary tumor, lung tumor, and erythroid cells. Table 1 also demonstrates that the expression of 25943 is upregulated in ovary tumors, as compared to normal ovary tissue; upregulated in colon tumors, as compared to normal colon tissue; and upregulated in lung tumors, as compared to lung tissue under conditions of chronic obstructive pulmonary disease (COPD).

EXAMPLE 9: EXPRESSION ANALYSIS OF HUMAN 25943 mRNA IN COLON CANCER USING TAQMAN AND IN SITU ANALYSIS

This example describes the expression of human 25943 mRNA in various colon tumors and cell lines, as determined using the TaqMan™ procedure (described above) and *in situ* hybridization analysis.

For *in situ* analysis, various tumors and normal tissues were first frozen on dry ice. Ten-micrometer-thick sections of the tissues were postfixed with 4% formaldehyde in DEPC-treated 1X phosphate-buffered saline at room temperature for 10 minutes before being rinsed twice in DEPC 1X phosphate-buffered saline and once in 0.1 M triethanolamine-HCl (pH 8.0). Following incubation in 0.25% acetic anhydride-0.1 M triethanolamine-HCl for 10 minutes, sections were rinsed in DEPC 2X SSC (1X SSC is 0.15 M NaCl plus 0.015 M sodium citrate). Tissue was then dehydrated through a series of ethanol washes, incubated in 100% chloroform for 5 minutes, and then rinsed in 100% ethanol for 1 minute and 95% ethanol for 1 minute and allowed to air dry.

Hybridizations were performed with ^{35}S-radiolabeled (5 X 10^7 cpm/ml) cRNA probes. Probes were incubated in the presence of a solution containing 600 mM NaCl, 10 mM Tris (pH 7.5), 1 mM EDTA, 0.01% sheared salmon sperm DNA, 0.01% yeast tRNA, 0.05% yeast total RNA type X1, 1X Denhardt's solution, 50% formamide, 10% dextran

sulfate, 100 mM dithiothreitol, 0.1% sodium dodecyl sulfate (SDS), and 0.1% sodium thiosulfate for 18 hours at 55°C.

After hybridization, slides were washed with 2X SSC. Sections were then sequentially incubated at 37°C in TNE (a solution containing 10 mM Tris-HCl (pH 7.6), 500 mM NaCl, and 1 mM EDTA), for 10 minutes, in TNE with 10μg of RNase A per ml for 30 minutes, and finally in TNE for 10 minutes. Slides were then rinsed with 2X SSC at room temperature, washed with 2X SSC at 50°C for 1 hour, washed with 0.2X SSC at 55° C for 1 hour, and 0.2X SSC at 60°C for 1 hour. Sections were then dehydrated rapidly through serial ethanol-0.3 M sodium acetate concentrations before being air dried and exposed to Kodak Biomax MR scientific imaging film for 24 hours and subsequently dipped in NB-2 photoemulsion and exposed at 4°C for 7 days before being developed and counter stained.

TaqMan - solid tumors

The expression of human 25943 was examined in solid human colon tumors. As shown in Table 2, 25943 expression is upregulated in 3/4 colon tumors, as compared to normal colon tissue. 25943 expression is also upregulated in a colon metastasis to the liver, as compared to normal liver tissue.

The expression human 25943 was also examined in various solid human colon tumor at different stages of tumorigenesis. As shown in Tables 5 and 6 (which are duplicate analyses of the same tumor samples), 25943 is highly expressed in 2/2 adenomas (as compared to normal colon), in 1/5 stage B adenocarcinomas (as compared to normal colon), in 2/6 stage C adenocarcinomas (as compared to normal colon), and in 1/5 liver metastases (as compared to normal liver and normal colon).

The expression human 25943 was further examined in colon metastases to the liver. As shown in Table 7, 25943 is highly expressed in 4/17 metastases, as compared to normal liver and normal colon. 25943 is also highly expressed in one early stage adenocarcinoma sample, as compared to normal colon.

In situ hybridization - solid tumors

In situ hybridization analysis indicated that human 25943 was expressed in tumors (3/3 positive samples) and metastatic carcinomas (3/4 positive samples), but not in

dysplastic/adenoma (0/3 positive samples), normal colonic epithelium (0/3 positive samples), or normal liver (0/2 positive samples).

TaqMan - in vitro models

The expression of human 25943 was examined in a number of xenograft friendly colon tumor cell lines. These cell lines can produce tumors when injected into mice. As shown in Tables 3 and 4, 25943 is highly expressed in DLD1 (stage C) cells, SW620 (stage C) cells, and HCT116 cells.

The expression of human 25943 was also examined in synchronized colon tumor cells induced to enter the cell cycle. As shown in Table 9, expression of 25943 was not noticeably regulated in HCT 116 cells synchronized with Aphidicolin, which blocks at the G1 stage of the cell cycle. Expression of 25943 was unregulated in HCT 116 cells during progression through the cell cycle after synchronization with Nocodazole, which blocks at the G2/M stage of the cell cycle. Expression of 25943 was also upregulated in DLD1 cells after synchronization with Nocodazole.

The expression of human 25943 was further examined in *in vitro* colon cancer models (Tables 8 and 10). Disruption of the k-ras gene in DLD1 colon cancer cells (samples 2-4, as compared to sample 1, in Table 8; samples 16-18, as compared to sample 14, in Table 10) and in HCT 116 cells (samples 6-9, as compared to sample 5, in Table 8; samples 19-22, as compared to sampled 15, in Table 10) resulted in the downregulation of the expression of 25943.

EXAMPLE 10: EXPRESSION ANALYSIS OF HUMAN 25943 mRNA IN OVARIAN CANCER USING TAQMAN AND IN SITU ANALYSIS

This example describes the expression of human 25943 mRNA in various ovarian tumors and cell lines, as determined using the TaqMan™ procedure and *in situ* hybridization analysis (as described above).

Solid tumors

The expression of human 25943 was examined in solid human ovarian tumors using Taqman analysis. As shown in Table 2, 25943 expression is strongly upregulated in 5/5 ovarian tumors, as compared to normal ovary.

The expression of human 25943 was examined in solid human ovarian tumors using *in situ* hybridization analysis. Strong 25943 expression was seen in all ovarian tumors types examined (5/7 tumors total), including adenocarcinoma, clear cell carcinoma, and serous carcinoma, as compared to normal ovary.

TaqMan - cell lines and tumor models

The expression of human 25943 was examined in various ovarian tumor models. As shown in Table 11, 25943 expression is downregulated in response to the growth factors EGF (samples 2-4, as compared to sample 1) and Heregulin (samples 4-6, as compared to sample 1) in the ovarian cancer cell line SKOV3 in the absence of serum. A cisplatin resistant SKOV-3 variant shows similar downregulation for EGF (samples 8 and 10-12), but not for Heregulin (samples 8 and 13-15). This data indicates that 25943 expression may be regulated through the epidermal growth factor receptor (EGFR) family, which includes EGFR, Her2, Her3 and Her4.

The expression of human 25943 was also examined in xenograft friendly ovarian tumor cell lines and other ovarian tumor cell lines. As shown in Tables 3 and 4, 25943 is highly expressed in SKOV-3 cells and OVCAR-3 cells. As shown in Table 11, 25943 is highly expressed in SKOV-3 cells, cisplatin resistant SKOV-3 variant cells, A2780 cells, OVCAR-3 cells, MDA2774 cells, and DOV13 cells.

The expression of human 25943 was further examined in serum treated HEY ovarian cancer cells. The cells were serum starved for 24 hours, and time points were taken at 0, 1, 3, 6, 9, and 12 hours after the addition of 10% serum. c-myc protein is highly upregulated at 1 hour after addition of serum, and phosphorylated at 6 hours. As shown in Table 11, samples 28-33, 25943 expression is upregulated between 1 and 3 hours after treatment, and then downregulated between 6 and 12 hours, indicating regulation by c-myc.

The expression of human 25943 was also examined in SKOV-3 and cisplatin resistant SKOV-3 variant subcutaneous xenograft tumors in nude mice. As shown in Table 11 (samples 34-37), 25943 expression is strongly downregulated in the tumors, as compared to the parental cells grown on plastic.

Finally, 25943 expression of human 25943 is upregulated in 1/2 ovarian ascites samples, as compared to normal ovary (Table 11, samples 38-41).

EXAMPLE 11: EXPRESSION ANALYSIS OF HUMAN 25943 mRNA IN LUNG CANCER USING TAQMAN AND *IN SITU* ANALYSIS

This example describes the expression of human 25943 mRNA in various lung tumors and cell lines, as determined using the TaqMan™ procedure and *in situ* hybridization analysis (as described above).

Solid tumors

The expression of human 25943 was examined in solid human lung tumors using Taqman analysis. As shown in Table 2, 25943 expression is upregulated in 4/6 lung tumors, as compared to normal lung. 25943 expression is also upregulated in a colon metastasis to the liver, as compared to normal liver.

The expression of human 25943 was examined in solid human lung tumors using *in situ* hybridization analysis. Moderate 25943 expression was seen 2/2 lung tumors, as compared to normal lung.

TaqMan - cell lines and tumor models

The expression of human 25943 was examined in xenograft friendly lung tumor cell lines. As shown in Tables 3 and 4, 25943 is highly expressed in NCIH125 cells, NCIH67 cells, A549 cells, and normal human bronchial epithelium cells (NHBE).

EXAMPLE 12: EXPRESSION ANALYSIS OF HUMAN 25943 mRNA IN BREAST CANCER USING TAQMAN AND *IN SITU* ANALYSIS

This example describes the expression of human 25943 mRNA in various breast tumors and cell lines, as determined using the TaqMan™ procedure and *in situ* hybridization analysis (as described above).

Solid tumors

The expression of human 25943 was examined in solid human breast tumors using *in situ* hybridization analysis. Moderate 25943 expression was seen 2/2 breast tumors (invasive ductal carcinoma), as compared to normal breast.

TaqMan - cell lines and tumor models

The expression of human 25943 was examined in xenograft friendly breast tumor cell lines. As shown in Tables 3 and 4, 25943 is highly expressed in MCF-7 cells, ZR75 cells, T47D cells, and SkBr3 cells.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

What is claimed:

1. A method for identifying a compound capable of treating a cell proliferation disorder, comprising assaying the ability of the compound to modulate 25943 nucleic acid expression or 25943 polypeptide activity, thereby identifying a compound capable of treating a cell proliferation disorder.

2. A method for identifying a compound capable of modulating cellular proliferation comprising:

a) contacting a cell which expresses 25943 with a test compound; and

b) assaying the ability of the test compound to modulate the expression of a 25943 nucleic acid or the activity of a 25943 polypeptide, thereby identifying a compound capable of modulating cellular proliferation.

3. A method for modulating cellular proliferation in a cell comprising contacting a cell with a 25943 modulator, thereby modulating cellular proliferation in the cell.

4. The method of claim 2, wherein the cell is a breast cell, an ovarian cell, a lung cell or a colon cell.

5. The method of claim 3, wherein the 25943 modulator is a small organic molecule, peptide, antibody or antisense nucleic acid molecule.

6. The method of claim 3, wherein the 25943 modulator is capable of modulating 25943 polypeptide activity.

7. The method of claim 6, wherein the 25943 modulator is a small organic molecule, peptide, antibody or antisense nucleic acid molecule.

8. The method of claim 6, wherein the 25943 modulator is capable of modulating 25943 nucleic acid expression.

9. A method for treating a subject having a cell proliferation disorder characterized by aberrant 25943 polypeptide activity or aberrant 25943 nucleic acid expression comprising administering to the subject a 25943 modulator, thereby treating said subject having a cell proliferation disorder.

10. The method of claim 9, wherein said cell proliferation disorder is selected from the group consisting of breast cancer, ovarian cancer, lung cancer and colon cancer.

11. The method of claim 9, wherein said 25943 modulator is administered in a pharmaceutically acceptable formulation.

12. The method of claim 9, wherein the 25943 modulator is a small organic molecule, peptide, antibody or antisense nucleic acid molecule.

13. The method of claim 9, wherein the 25943 modulator is capable of modulating 25943 polypeptide activity.

SEQUENCE LISTING

<110> Williamson, Mark
Rudolph-Owen, Laura A.

<120> METHODS AND COMPOSITIONS FOR THE
TREATMENT AND DIAGNOSIS OF CELLULAR PROLIFERATION DISORDERS
USING 25943

<130> MPI01-236

<150> 60/335,004
<151> 2001-10-31

<160> 3

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 1361
<212> DNA
<213> Homo sapien

<400> 1
```
gggcgggctg agcggtttcg agccggcgtc ggggagcggc ggtaccgggc ggctgcgggg 60
ctggctcgac ccagcttgag gtctcggcgt ccgcgtcctg cggtgccctg ggatccgccg 120
acatgaatcc catcgtagtg gtccacggcg gcggagccgg tcccatctcc aaggatcgga 180
aggagcgagt gcaccagggc atggtcagag ccgccaccgt gggctacggc atcctccggg 240
agggcgggag cgccgtggat gccgtagagg gagctgtcgt cgccctggaa gacgatcccg 300
agttcaacgc aggttgtggg tctgtcttga acacaaatgg tgaggttgaa atggatgcta 360
gtatcatgga tggaaaagac ctgtctgcag gagcagtgtc cgcagtccag tgtatagcaa 420
atcccattaa acttgctcgg cttgtcatgg aaaagacacc tcattgcttt ctgactgacc 480
aaggcgcagc gcagtttgca gcagctatgg gggttccaga gattcctgga gaaaaactgg 540
tgacagagag aaacaaaaag cgcctggaaa aagagaagca tgaaaaaggt gctcagaaaa 600
cagattgtca aaaaaacttg ggaaccgtgg gtgctgttgc cttggactgc aaagggaatg 660
tagcctacgc aacctccaca ggcggtatcg ttaataaaat ggtcggccgc gttggggact 720
caccgtgtct aggagctgga ggttatgccg acaatgacat cggagccgtc tcaaccacag 780
ggcatgggga aagcatcctg aaggtgaacc tggctagact caccctgttc cacatagaac 840
aaggaaagac ggtagaagag gctgcggacc tatcgttggg ttatatgaag tcaagggtta 900
aaggtttagg tggcctcatc gtggttagca aaacaggaga ctgggtggca aagtggacct 960
ccacctccat gccctgggca gccgccaagg acggcaagct gcacttcgga attgatcctg 1020
acgatactac tatcaccgac cttccctaag ccgctggaag attgtattcc agatgctagc 1080
ttagaggtca agtacagtct cctcatgaga catagcctaa tcaattagat ctagaattgg 1140
aaaaattgtc ccgtctgtca cttgttttgt tgccttaata agcatctgaa tgtttggttg 1200
tggggcgggt tctgaagcga tgagagaaat gcccgtatta ggaggattac ttgagccctg 1260
gaggtcaaag ctgaggtgag ccatgattac tccactgcac tccagcctgg gcaacagagc 1320
caggccctgt atcaaaaaaa aaaaaaaaaa aaaaaaagaa a                    1361
```

<210> 2
<211> 308
<212> PRT
<213> Homo sapien

<400> 2
```
Met Asn Pro Ile Val Val Val His Gly Gly Gly Ala Gly Pro Ile Ser
 1               5                  10                  15
Lys Asp Arg Lys Glu Arg Val His Gln Gly Met Val Arg Ala Ala Thr
            20                  25                  30
Val Gly Tyr Gly Ile Leu Arg Glu Gly Gly Ser Ala Val Asp Ala Val
        35                  40                  45
Glu Gly Ala Val Val Ala Leu Glu Asp Asp Pro Glu Phe Asn Ala Gly
```

```
    50                  55                  60
Cys Gly Ser Val Leu Asn Thr Asn Gly Glu Val Glu Met Asp Ala Ser
65                  70                  75                  80
Ile Met Asp Gly Lys Asp Leu Ser Ala Gly Ala Val Ser Ala Val Gln
                85                  90                  95
Cys Ile Ala Asn Pro Ile Lys Leu Ala Arg Leu Val Met Glu Lys Thr
            100                 105                 110
Pro His Cys Phe Leu Thr Asp Gln Gly Ala Ala Gln Phe Ala Ala Ala
        115                 120                 125
Met Gly Val Pro Glu Ile Pro Gly Glu Lys Leu Val Thr Glu Arg Asn
    130                 135                 140
Lys Lys Arg Leu Glu Lys Glu Lys His Glu Lys Gly Ala Gln Lys Thr
145                 150                 155                 160
Asp Cys Gln Lys Asn Leu Gly Thr Val Gly Ala Val Ala Leu Asp Cys
                165                 170                 175
Lys Gly Asn Val Ala Tyr Ala Thr Ser Thr Gly Gly Ile Val Asn Lys
            180                 185                 190
Met Val Gly Arg Val Gly Asp Ser Pro Cys Leu Gly Ala Gly Gly Tyr
        195                 200                 205
Ala Asp Asn Asp Ile Gly Ala Val Ser Thr Thr Gly His Gly Glu Ser
    210                 215                 220
Ile Leu Lys Val Asn Leu Ala Arg Leu Thr Leu Phe His Ile Glu Gln
225                 230                 235                 240
Gly Lys Thr Val Glu Glu Ala Ala Asp Leu Ser Leu Gly Tyr Met Lys
                245                 250                 255
Ser Arg Val Lys Gly Leu Gly Gly Leu Ile Val Val Ser Lys Thr Gly
            260                 265                 270
Asp Trp Val Ala Lys Trp Thr Ser Thr Ser Met Pro Trp Ala Ala Ala
        275                 280                 285
Lys Asp Gly Lys Leu His Phe Gly Ile Asp Pro Asp Asp Thr Thr Ile
    290                 295                 300
Thr Asp Leu Pro
305
```

<210> 3
<211> 1361
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (123)...(1049)

<400> 3

```
gggcgggctg agcggtttcg agccggcgtc ggggagcggc ggtaccgggc ggctgcgggg 60
ctggctcgac ccagcttgag gtctcggcgt ccgcgtcctg cggtgccctg ggatccgccg 120
ac atg aat ccc atc gta gtg gtc cac ggc ggc gga gcc ggt ccc atc    167
   Met Asn Pro Ile Val Val Val His Gly Gly Gly Ala Gly Pro Ile
    1               5                   10                  15

tcc aag gat cgg aag gag cga gtg cac cag ggc atg gtc aga gcc gcc   215
Ser Lys Asp Arg Lys Glu Arg Val His Gln Gly Met Val Arg Ala Ala
                20                  25                  30

acc gtg ggc tac ggc atc ctc cgg gag ggc ggg agc gcc gtg gat gcc   263
Thr Val Gly Tyr Gly Ile Leu Arg Glu Gly Gly Ser Ala Val Asp Ala
            35                  40                  45

gta gag gga gct gtc gtc gcc ctg gaa gac gat ccc gag ttc aac gca   311
Val Glu Gly Ala Val Val Ala Leu Glu Asp Asp Pro Glu Phe Asn Ala
```

50 55 60

ggt tgt ggg tct gtc ttg aac aca aat ggt gag gtt gaa atg gat gct 359
Gly Cys Gly Ser Val Leu Asn Thr Asn Gly Glu Val Glu Met Asp Ala
65 70 75

agt atc atg gat gga aaa gac ctg tct gca gga gca gtg tcc gca gtc 407
Ser Ile Met Asp Gly Lys Asp Leu Ser Ala Gly Ala Val Ser Ala Val
80 85 90 95

cag tgt ata gca aat ccc att aaa ctt gct cgg ctt gtc atg gaa aag 455
Gln Cys Ile Ala Asn Pro Ile Lys Leu Ala Arg Leu Val Met Glu Lys
100 105 110

aca cct cat tgc ttt ctg act gac caa ggc gca gcg cag ttt gca gca 503
Thr Pro His Cys Phe Leu Thr Asp Gln Gly Ala Ala Gln Phe Ala Ala
115 120 125

gct atg ggg gtt cca gag att cct gga gaa aaa ctg gtg aca gag aga 551
Ala Met Gly Val Pro Glu Ile Pro Gly Glu Lys Leu Val Thr Glu Arg
130 135 140

aac aaa aag cgc ctg gaa aaa gag aag cat gaa aaa ggt gct cag aaa 599
Asn Lys Lys Arg Leu Glu Lys Glu Lys His Glu Lys Gly Ala Gln Lys
145 150 155

aca gat tgt caa aaa aac ttg gga acc gtg ggt gct gtt gcc ttg gac 647
Thr Asp Cys Gln Lys Asn Leu Gly Thr Val Gly Ala Val Ala Leu Asp
160 165 170 175

tgc aaa ggg aat gta gcc tac gca acc tcc aca ggc ggt atc gtt aat 695
Cys Lys Gly Asn Val Ala Tyr Ala Thr Ser Thr Gly Gly Ile Val Asn
180 185 190

aaa atg gtc ggc cgc gtt ggg gac tca ccg tgt cta gga gct gga ggt 743
Lys Met Val Gly Arg Val Gly Asp Ser Pro Cys Leu Gly Ala Gly Gly
195 200 205

tat gcc gac aat gac atc gga gcc gtc tca acc aca ggg cat ggg gaa 791
Tyr Ala Asp Asn Asp Ile Gly Ala Val Ser Thr Thr Gly His Gly Glu
210 215 220

agc atc ctg aag gtg aac ctg gct aga ctc acc ctg ttc cac ata gaa 839
Ser Ile Leu Lys Val Asn Leu Ala Arg Leu Thr Leu Phe His Ile Glu
225 230 235

caa gga aag acg gta gaa gag gct gcg gac cta tcg ttg ggt tat atg 887
Gln Gly Lys Thr Val Glu Glu Ala Ala Asp Leu Ser Leu Gly Tyr Met
240 245 250 255

aag tca agg gtt aaa ggt tta ggt ggc ctc atc gtg gtt agc aaa aca 935
Lys Ser Arg Val Lys Gly Leu Gly Gly Leu Ile Val Val Ser Lys Thr
260 265 270

gga gac tgg gtg gca aag tgg acc tcc acc tcc atg ccc tgg gca gcc 983
Gly Asp Trp Val Ala Lys Trp Thr Ser Thr Ser Met Pro Trp Ala Ala
275 280 285

gcc aag gac ggc aag ctg cac ttc gga att gat cct gac gat act act 1031
Ala Lys Asp Gly Lys Leu His Phe Gly Ile Asp Pro Asp Asp Thr Thr
290 295 300

```
atc acc gac ctt ccc taa gccgctggaa gattgtattc cagatgctag          1079
Ile Thr Asp Leu Pro  *
    305

cttagaggtc aagtacagtc tcctcatgag acatagccta atcaattaga tctagaattg 1139
gaaaaattgt cccgtctgtc acttgttttg ttgccttaat aagcatctga atgtttggtt 1199
gtggggcggg ttctgaagcg atgagagaaa tgcccgtatt aggaggatta cttgagccct 1259
ggaggtcaaa gctgaggtga gccatgatta ctccactgca ctccagcctg ggcaacagag 1319
ccaggccctg tatcaaaaaa aaaaaaaaaa aaaaaaaaga aa                    1361
```